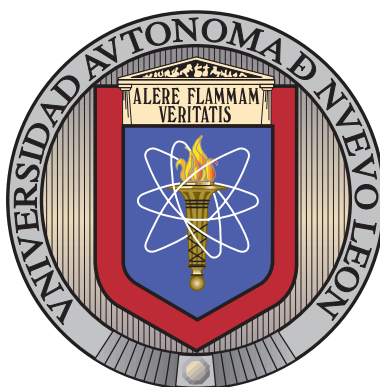


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO



MODELOS ESTADÍSTICO-ESPACIALES DE
CONTAMINANTES DEL AIRE EN EL ÁREA
METROPOLITANA DE MONTERREY

POR

JORGE ARMANDO SERNA MENDOZA

COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS DE LA INGENIERÍA
CON ORIENTACIÓN EN SISTEMAS

NOVIEMBRE 2020

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO



MODELOS ESTADÍSTICO-ESPACIALES DE
CONTAMINANTES DEL AIRE EN EL ÁREA
METROPOLITANA DE MONTERREY

POR

JORGE ARMANDO SERNA MENDOZA

COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS DE LA INGENIERÍA
CON ORIENTACIÓN EN SISTEMAS

NOVIEMBRE 2020



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica
Subdirección de Estudios de Posgrado

Los miembros del Comité de Tesis recomendamos que la Tesis «Modelos estadístico-espaciales de contaminantes del aire en el Área Metropolitana de Monterrey», realizada por el alumno Jorge Armando Serna Mendoza, con número de matrícula 1985281, sea aceptada para su defensa como requisito para obtener el grado de Maestría en Ciencias de la Ingeniería con Orientación en Sistemas.

El Comité de Tesis

Dr. José Arturo Berrones Santos
Director

Dra. Satu Elisa Schaeffer
Revisora

Dra. Bibiana Obregón Quintana
Revisora

Vo. Bo

Dr. Simón Martínez Martínez
Subdirector de Estudios de Posgrado



FIME
069

San Nicolás de los Garza, Nuevo León, octubre de 2020

*A mis padres, hermanas, novia y
a esa persona que esta en el cielo,
mandandome salud porque eso es lo primero.*

ÍNDICE GENERAL

Agradecimientos	xxiv
1. Resumen	1
2. Introducción	3
2.1. Panorama General	3
2.1.1. Definición del Problema	5
2.1.2. Objetivo	6
2.1.3. Hipótesis	6
2.1.4. Alcance	7
2.1.5. Motivación	7
2.1.6. Justificación	9
2.2. Estructura de la Tesis	9

3. Antecedentes y Aplicaciones	11
3.1. Calidad del Aire en el AMM	11
3.2. Métodos de Interpolación	26
3.2.1. Teselaciones de Voronoi	26
3.2.2. Funciones de Base Radial	28
3.2.3. Distancia Inversa Ponderada	30
3.2.4. Kriging Ordinario	32
3.2.5. Kriging Universal	34
4. Marco Teórico	36
4.1. Análisis Estadístico	36
4.1.1. Series de Tiempo	37
4.1.2. Estacionariedad	38
4.1.3. Autocovarianzas y Autocorrelaciones	40
4.1.4. Variograma	42
4.2. Métodos de Interpolación	43
4.2.1. Determininistas	43
4.2.2. Geoestadísticos	47

5. Metodología y Experimentación	51
5.1. Análisis Estadístico y Limpieza de la Base de Datos	51
5.1.1. Monóxido de Carbono (CO)	57
5.1.2. Óxido Nítrico (NO)	61
5.1.3. Bióxido de Nitrógeno (NO ₂)	65
5.1.4. Óxidos de Nitrógeno (NO _X)	69
5.1.5. Ozono (O ₃)	73
5.1.6. Partículas Menores a 10 Micras (PM ₁₀)	77
5.1.7. Partículas Menores a 2.5 Micras (PM _{2.5})	81
5.1.8. Presión Barométrica	85
5.1.9. Precipitación Pluvial	89
5.1.10. Humedad Relativa	93
5.1.11. Bióxido de Azufre (SO ₂)	97
5.1.12. Radiación Solar	101
5.1.13. Temperatura Ambiental	105
5.1.14. Velocidad del Viento	109
5.1.15. Dirección del Viento	113

5.2. Medición del Error	116
5.2.1. Error	117
5.2.2. MAPE	118
5.2.3. MAE	118
5.2.4. RMSE	119
5.2.5. MSE	119
5.3. Entorno Computacional y Tiempo de Ejecución	120
6. Resultados y Conclusiones	123
6.1. Variables del <i>Índice de AIRE y SALUD</i>	125
6.1.1. Monóxido de Carbono (CO)	125
6.1.2. Dióxido de Nitrógeno (NO ₂)	133
6.1.3. Ozono (O ₃)	141
6.1.4. Partículas Menores a 10 Micras (PM ₁₀)	149
6.1.5. Partículas Menores a 2.5 Micras (PM _{2.5})	157
6.1.6. Bióxido de Azufre (SO ₂)	165
6.2. Resto de Variables	173
6.2.1. Óxido Nítrico (NO)	173

6.2.2. Óxidos de Nitrógeno (NO_X)	182
6.2.3. Velocidad del Viento	190
6.2.4. Dirección del Viento	198
6.3. Conclusiones	206
6.3.1. Monóxido de Carbono (CO)	206
6.3.2. Bióxido de Nitrógeno (NO_2)	206
6.3.3. Ozono (O_3)	207
6.3.4. Partículas Menores a 10 Micras (PM_{10})	207
6.3.5. Partículas Menores a 2.5 Micras ($\text{PM}_{2.5}$)	208
6.3.6. Bióxido de Azufre (SO_2)	208
6.3.7. Óxido Nítrico (NO)	208
6.3.8. Óxidos de Nitrógeno (NO_X)	209
6.3.9. Velocidad del Viento	209
6.3.10. Dirección del Viento	210
6.4. Trabajo Futuro	210

ÍNDICE DE FIGURAS

3.1. Red de monitoreo del aire localizado en Monterrey N.L.	14
3.2. Mapa del AMM con las trece estaciones meteorológicas marcadas por un símbolo de posición geográfica (Secretaría de Desarrollo Sustenta- ble de Nuevo León, 2020)	22
4.1. Gráfico de la serie de tiempo del rendimiento del mercado de valores del tesoro de EE. UU. con vencimiento constante a 10 años (Montgo- mery <i>et al.</i> , 2015)	37
4.2. Series de tiempo	39
4.3. Datos correlacionados y no correlacionados	40
5.1. Datos conocidos y desconocidos (NaN) registrados por SIMA	54
5.2. Series de tiempo de CO desde 2016 hasta 2018	57
5.3. Histogramas de CO	58
5.4. Matriz de correlación de CO entre estaciones	59

5.5. Matriz de correlación de CO contra el resto de las variables	60
5.6. Series de tiempo de NO desde 2016 hasta 2018	61
5.7. Histogramas de NO	62
5.8. Matriz de correlación de NO entre estaciones	63
5.9. Matriz de correlación de NO contra el resto de las variables	64
5.10. Series de tiempo de NO ₂ desde 2016 hasta 2018	65
5.11. Histogramas de NO ₂	66
5.12. Matriz de correlación de NO ₂ entre estaciones	67
5.13. Matriz de correlación de NO ₂ contra el resto de las variables	68
5.14. Series de tiempo de NO _X desde 2016 hasta 2018	69
5.15. Histogramas de NO _X	70
5.16. Matriz de correlación de NO _X entre estaciones	71
5.17. Matriz de correlación de NO _X contra el resto de las variables	72
5.18. Series de tiempo de O ₃ desde 2016 hasta 2018	73
5.19. Histogramas de O ₃	74
5.20. Matriz de correlación de O ₃ entre estaciones	75
5.21. Matriz de correlación de O ₃ contra el resto de las variables	76

5.22. Series de tiempo de PM_{10} desde 2016 hasta 2018	77
5.23. Histogramas de PM_{10}	78
5.24. Matriz de correlación de PM_{10} entre estaciones	79
5.25. Matriz de correlación de PM_{10} contra el resto de las variables	80
5.26. Series de tiempo de $PM_{2.5}$ desde 2016 hasta 2018	81
5.27. Histogramas de $PM_{2.5}$	82
5.28. Matriz de correlación de $PM_{2.5}$ entre estaciones	83
5.29. Matriz de correlación de $PM_{2.5}$ contra el resto de las variables	84
5.30. Series de tiempo de <i>presión barométrica</i> desde 2016 hasta 2018	85
5.31. Histogramas de <i>presión barométrica</i>	86
5.32. Matriz de correlación de <i>presión barométrica</i> entre estaciones	87
5.33. Matriz de correlación de <i>presión barométrica</i> contra el resto de las variables	88
5.34. Series de tiempo de <i>precipitación pluvial</i> desde 2016 hasta 2018	89
5.35. Histogramas de <i>precipitación pluvial</i>	90
5.36. Matriz de correlación de <i>precipitación pluvial</i> entre estaciones	91
5.37. Matriz de correlación de <i>precipitación pluvial</i> contra el resto de las variables	92

5.38. Series de tiempo de <i>humedad relativa</i> desde 2016 hasta 2018	93
5.39. Histogramas de <i>humedad relativa</i>	94
5.40. Matriz de correlación de <i>humedad relativa</i> entre estaciones	95
5.41. Matriz de correlación de <i>humedad relativa</i> contra el resto de las variables	96
5.42. Series de tiempo de SO ₂ desde 2016 hasta 2018	97
5.43. Histogramas de SO ₂	98
5.44. Matriz de correlación de SO ₂ entre estaciones	99
5.45. Matriz de correlación SO ₂ contra el resto de las variables	100
5.46. Series de tiempo de <i>radiación solar</i> desde 2016 hasta 2018	101
5.47. Histogramas de <i>radiación solar</i>	102
5.48. Matriz de correlación de <i>radiación solar</i> entre estaciones	103
5.49. Matriz de correlación de <i>radiación solar</i> contra el resto de las variables	104
5.50. Series de tiempo de <i>temperatura ambiental</i> desde 2016 hasta 2018 . .	105
5.51. Histogramas de <i>temperatura ambiental</i>	106
5.52. Matriz de correlación de <i>temperatura ambiental</i> entre estaciones . . .	107
5.53. Matriz de correlación de <i>temperatura ambiental</i> contra el resto de las variables	108

5.54. Series de tiempo de <i>velocidad del viento</i> desde 2016 hasta 2018	109
5.55. Histogramas de <i>velocidad del viento</i>	110
5.56. Matriz de correlación de <i>velocidad del viento</i> entre estaciones	111
5.57. Matriz de correlación de <i>velocidad del viento</i> contra el resto de las variables	112
5.58. Series de tiempo de <i>dirección del viento</i> desde 2016 hasta 2018	113
5.59. Histogramas de <i>dirección del viento</i>	114
5.60. Matriz de correlación de <i>dirección del viento</i> entre estaciones	115
5.61. Matriz de correlación de <i>dirección del viento</i> contra el resto de las variables	116
5.62. Tiempos de ejecución	121
6.1. Interpolaciones de CO para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	126
6.2. Interpolaciones de CO para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	128
6.3. Interpolaciones de CO para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	130
6.4. Interpolaciones de CO para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	132

6.5. Interpolaciones de NO ₂ para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	134
6.6. Interpolaciones de NO ₂ para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	136
6.7. Interpolaciones de NO ₂ para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	138
6.8. Interpolaciones de NO ₂ para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	140
6.9. Interpolaciones de O ₃ para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	142
6.10. Interpolaciones de O ₃ para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	144
6.11. Interpolaciones de O ₃ para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	146
6.12. Interpolaciones de O ₃ para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	148
6.13. Interpolaciones de PM ₁₀ para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	150
6.14. Interpolaciones de PM ₁₀ para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	152
6.15. Interpolaciones de PM ₁₀ para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	154

6.16. Interpolaciones de PM_{10} para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	156
6.17. Interpolaciones de $PM_{2.5}$ para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	158
6.18. Interpolaciones de $PM_{2.5}$ para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	160
6.19. Interpolaciones de $PM_{2.5}$ para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	162
6.20. Interpolaciones de $PM_{2.5}$ para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	164
6.21. Interpolaciones de SO_2 para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	166
6.22. Interpolaciones de SO_2 para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	168
6.23. Interpolaciones de SO_2 para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	170
6.24. Interpolaciones de SO_2 para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	172
6.25. Interpolaciones de NO para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	175
6.26. Interpolaciones de NO para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	177

6.27. Interpolaciones de NO para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	179
6.28. Interpolaciones de NO para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	181
6.29. Interpolaciones de NO _x para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	183
6.30. Interpolaciones de NO _x para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	185
6.31. Interpolaciones de NO _x para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	187
6.32. Interpolaciones de NO _x para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	189
6.33. Interpolaciones de <i>velocidad del viento</i> para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	191
6.34. Interpolaciones de <i>velocidad del viento</i> para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	193
6.35. Interpolaciones de <i>velocidad del viento</i> para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	195
6.36. Interpolaciones de <i>velocidad del viento</i> para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	197
6.37. Interpolaciones de <i>dirección del viento</i> para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	199

6.38. Interpolaciones de <i>dirección del viento</i> para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	201
6.39. Interpolaciones de <i>dirección del viento</i> para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)	203
6.40. Interpolaciones de <i>dirección del viento</i> para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)	205
6.41. Interpolaciones de CO	212
6.42. Interpolaciones de NO ₂	213
6.43. Interpolaciones de O ₃	214
6.44. Interpolaciones de PM ₁₀	215
6.45. Interpolaciones de PM _{2.5}	216
6.46. Interpolaciones de SO ₂	217
6.47. Interpolaciones de NO	218
6.48. Interpolaciones de NO _x	219

ÍNDICE DE TABLAS

3.1. Cumplimiento a las Normas de Calidad de Aire en la AMM (1993–2000)	15
3.2. Días Arriba de la Norma en la AMM (1993–2000)	16
3.3. Cumplimiento Anual de IMECAs Máximos en la AMM (1993–2000) .	17
3.4. Regiones con Mayores IMECAs Máximos en la AMM (1993–2000) . .	17
3.5. Cumplimiento Mensual de la Calidad del Aire en la AMM (1993–2000)	18
3.6. Comportamiento Hora a Hora de la Calidad del Aire en la AMM (1993–2000)	18
3.7. Características de las estaciones de monitoreo del SIMA	21
3.8. Variables y contaminantes medidos por SIMA	22
3.9. Concentraciones base para el cálculo del <i>Índice AIRE y SALUD</i> para cada contaminante	23
3.10. Obtención del <i>Índice AIRE y SALUD</i> para PM ₁₀	24
3.11. Obtención del <i>Índice AIRE y SALUD</i> para PM _{2.5}	24

3.12. Obtención del <i>Índice AIRE y SALUD</i> para O ₃	24
3.13. Obtención del <i>Índice AIRE y SALUD</i> para NO ₂	25
3.14. Obtención del <i>Índice AIRE y SALUD</i> para SO ₂	25
3.15. Obtención del <i>Índice AIRE y SALUD</i> para CO	25
3.16. Categorías del <i>Índice AIRE y SALUD</i>	26
5.1. Contaminantes y variables que recopilan las estaciones de monitoreo del SIMA con los nombres correspondientes de las variables en el código fuente (Serna, 2020)	53
5.2. Valores fuera de rango por falla en analizadores y sensores de los diferentes parámetros en las estaciones de monitoreo atmosférico . . .	55
6.1. CO: 9 estaciones seleccionadas 4 estaciones interpoladas	125
6.2. CO: 10 estaciones seleccionadas 3 estaciones interpoladas	127
6.3. CO: 11 estaciones seleccionadas 2 estaciones interpoladas	129
6.4. CO: 12 estaciones seleccionadas 1 estación interpolada	131
6.5. NO ₂ : 9 estaciones seleccionadas 4 estaciones interpoladas	133
6.6. NO ₂ : 10 estaciones seleccionadas 3 estaciones interpoladas	135
6.7. NO ₂ : 11 estaciones seleccionadas 2 estaciones interpoladas	137
6.8. NO ₂ : 12 estaciones seleccionadas 1 estación interpolada	139

6.9. O ₃ : 9 estaciones seleccionadas 4 estaciones interpoladas	141
6.10. O ₃ : 10 estaciones seleccionadas 3 estaciones interpoladas	143
6.11. O ₃ : 11 estaciones seleccionadas 2 estaciones interpoladas	145
6.12. O ₃ : 12 estaciones seleccionadas 1 estación interpolada	147
6.13. PM ₁₀ : 9 estaciones seleccionadas 4 estaciones interpoladas	149
6.14. PM ₁₀ : 10 estaciones seleccionadas 3 estaciones interpoladas	151
6.15. PM ₁₀ : 11 estaciones seleccionadas 2 estaciones interpoladas	153
6.16. PM ₁₀ : 12 estaciones seleccionadas 1 estación interpolada	155
6.17. PM _{2.5} : 9 estaciones seleccionadas 4 estaciones interpoladas	157
6.18. PM _{2.5} : 10 estaciones seleccionadas 3 estaciones interpoladas	159
6.19. PM _{2.5} : 11 estaciones seleccionadas 2 estaciones interpoladas	161
6.20. PM _{2.5} : 12 estaciones seleccionadas 1 estación interpolada	163
6.21. SO ₂ : 9 estaciones seleccionadas 4 estaciones interpoladas	165
6.22. SO ₂ : 10 estaciones seleccionadas 3 estaciones interpoladas	167
6.23. SO ₂ : 11 estaciones seleccionadas 2 estaciones interpoladas	169
6.24. SO ₂ : 12 estaciones seleccionadas 1 estación interpolada	171
6.25. NO: 9 estaciones seleccionadas 4 estaciones interpoladas	173

6.26. NO: 10 estaciones seleccionadas 3 estaciones interpoladas	176
6.27. NO: 11 estaciones seleccionadas 2 estaciones interpoladas	178
6.28. NO: 12 estaciones seleccionadas 1 estación interpolada	180
6.29. NO _X : 9 estaciones seleccionadas 4 estaciones interpoladas	182
6.30. NO _X : 10 estaciones seleccionadas 3 estaciones interpoladas	184
6.31. NO _X : 11 estaciones seleccionadas 2 estaciones interpoladas	186
6.32. NO _X : 12 estaciones seleccionadas 1 estación interpolada	188
6.33. <i>Velocidad del viento</i> : 9 estaciones seleccionadas 4 estaciones interpoladas	190
6.34. <i>Velocidad del viento</i> : 10 estaciones seleccionadas 3 estaciones interpo- ladas	192
6.35. <i>Velocidad del viento</i> : 11 estaciones seleccionadas 2 estaciones interpo- ladas	194
6.36. <i>Velocidad del viento</i> : 12 estaciones seleccionadas 1 estación interpolada	196
6.37. <i>Dirección del viento</i> : 9 estaciones seleccionadas 4 estaciones interpoladas	198
6.38. <i>Dirección del viento</i> : 10 estaciones seleccionadas 3 estaciones interpo- ladas	200
6.39. <i>Dirección del viento</i> : 11 estaciones seleccionadas 2 estaciones interpo- ladas	202

6.40. <i>Dirección del viento</i> : 12 estaciones seleccionadas 1 estaciones interpo-	
ladas	204

AGRADECIMIENTOS

Deseo agradecer a la Universidad Autónoma de Nuevo León (UANL) la oportunidad que me ha brindado de realizar mis estudios de posgrado. A la Facultad de Ingeniería Mecánica y Eléctrica (FIME) por formar parte de su alumnado. Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado mediante una beca de estudios de tiempo completo.

Agradezco al Posgrado en Ingeniería de Sistemas (PISIS) por ofrecerme la oportunidad de realizar mis estudios de maestría, en especial a mi asesor Arturo Berrones, quien me orientó y apoyó completamente en este proceso, y a mi revisora Elisa Schaeffer, quien me compartió sus conocimientos y por formar parte del comité de este trabajo de tesis. También, a mi revisora Bibiana Obregón, quien me ha ayudado a construir mi camino del saber y por formar parte del comité de este trabajo de tesis.

Agradezco los conocimientos brindados por mis profesores del posgrado, los comentarios y aportaciones de mis compañeros de generación y de otras generaciones, y también, agradezco las tazas de café y desvelos que compartí con mis amigos.

CAPÍTULO 1

RESUMEN

Jorge Armando Serna Mendoza.

Candidato para obtener el grado de Maestría en Ciencias de la Ingeniería con Orientación en Sistemas.

Universidad Autónoma de Nuevo León.

Facultad de Ingeniería Mecánica y Eléctrica.

Título del estudio: MODELOS ESTADÍSTICO-ESPACIALES DE CONTAMINANTES DEL AIRE EN EL ÁREA METROPOLITANA DE MONTERREY.

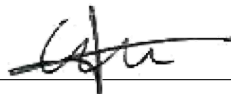
Número de páginas: 214.

OBJETIVOS Y MÉTODO DE ESTUDIO: Encontrar entre los métodos existentes de interpolación espacial, el método que mejor pronostique niveles de contaminación del aire en el Área Metropolitana de Monterrey. Los métodos seleccionados son: Teselaciones de Voronoi, Distancia Inversa Ponderada, Funciones de Base Radial y Kriging, los cuales se utilizan para pronosticar las variables

Ozono, Monóxido de Carbono, Óxido Nítrico, Bióxido de Nitrógeno, Óxidos de Nitrógeno, Partículas menores a 10 micras, Partículas menores a 2.5 micras y Bióxido de Azufre. Dichos datos de las variables son proporcionados por el Sistema Integral de Monitoreo Ambiental de la Secretaría de Desarrollo Sustentable del Gobierno de Nuevo León.

CONTRIBUCIONES Y CONCLUSIONES: Generar mapas que describan la calidad del aire sobre toda la área de estudio para cada contaminante, ya que cada variable sigue una distribución de los datos diferente a la del resto y así evaluar el Índice de AIRE y SALUD propuesto en la Norma Oficial Mexicana NOM-020-SSA1-2014, el cual indica los riesgos asociados a los niveles de contaminación de cada contaminante del aire.

Firma del asesor: _____



Dr. José Arturo Berrones Santos

CAPÍTULO 2

INTRODUCCIÓN

En este capítulo se presenta la introducción de este trabajo de tesis, donde se menciona el panorama general del problema de contaminación del aire en la Área Metropolitana de Monterrey (AMM) y cómo esta mala calidad del aire influye en la calidad de vida de sus residentes. Luego, continúa con la definición del problema, donde se delimita el área de estudio de este trabajo; seguido del objetivo, las hipótesis, alcance, motivación y justificación del presente trabajo.

2.1. Panorama General

Las ciudades o zonas metropolitanas son por naturaleza agrupamientos de personas y actividades, que por lo general son los lugares del planeta que presentan los niveles más altos de contaminación (Fenger, 1999). Las actividades diarias en zonas metropolitanas generan una gran cantidad de sustancias contaminantes que modifican la composición del aire, sumado a esto, la dirección del viento propicia el esparcimiento e interacción de sustancias entre distintos lugares, es decir, las sustancias presentes en un determinado lugar pueden haber sido generadas ahí o pudieron haber llegado de algún de otro lugar (Mayer,

1999).

La presencia de dichas sustancias contaminantes provoca el deterioro de la calidad del aire, el cual propicia efectos dañinos en el medio ambiente y en la salud humana. Actualmente (octubre de 2020) el Área Metropolitana de Monterrey (AMM) presenta problemas de calidad del aire, debido al aumento de las actividades dentro de ésta, como son los centros de trabajo, centros de estudio, centros comerciales, entre otros.

El AMM cuenta con trece estaciones fijas de monitoreo bajo el encargo del Sistema Integral de Monitoreo Ambiental de la Secretaría de Desarrollo Sustentable del Gobierno de Nuevo León (SIMA), las cuales algunas de ellas han estado en funcionamiento desde el año 1993. Los datos registrados por las estaciones de monitoreo dan una indicación de las tendencias en la calidad del aire en el AMM, sin embargo, las estaciones de monitoreo están ubicadas en sitios fijos, por lo que la información reportada solo representa microambientes alrededor de ellas.

Los estudios de evaluación de la calidad del aire tienen altos costos de muestreo o análisis, ya que se requiere de equipos de alta tecnología. Además, el representar el comportamiento de la calidad del aire con pocos puntos es un problema retante y los datos resultan ser falsos cuando se extrapolan sobre toda la región, especialmente en las zonas metropolitanas de países en desarrollo (Bayraktar y Turalioglu, 2005).

En este trabajo, las mediciones de los contaminantes del aire de las trece estaciones de monitoreo se consideran para la determinación de puntos espaciales críticos sobre el AMM, a través de métodos de interpolación espacial deterministas y probabilísticos, para predecir valores desconocidos en puntos no muestreados en el AMM, a partir de los valores conocidos de los puntos muestreados en las trece estaciones de monitoreo.

Diferentes métodos pueden producir resultados de puntos interpolados distintos (Tang *et al.*, 2016), por lo que se requiere una comprensión completa del problema para encontrar métodos de interpolación adecuados para la región del AMM.

2.1.1 DEFINICIÓN DEL PROBLEMA

En el área de gobernación, la obligación de informar a la población sobre la calidad del aire donde residen sus habitantes es importante para establecer medidas preventivas que cuiden la salud de los habitantes. Además, la necesidad de conocer las mediciones históricas en una ciudad son importantes para verificar y actualizar la efectividad de los programas implementados para el control o reducción de la contaminación urbana. Sumado a esto, en el área científica los datos de mediciones de aire son componentes primordiales para entender, modelar, simular, interpolar, pronosticar y desarrollar modelos que encaminen a la creación de programas para el control de la contaminación atmosférica (Martínez *et al.*, 2001).

Resultados de estudios científicos internacionales de la salud pública han propiciado a la creación de las Normas Oficiales Mexicanas (NOM), las cuales definen las concentraciones y los tiempos de exposición en los que un ciudadano pueda desenvolverse sin que su salud se vea afectada (Martínez *et al.*, 2001). Dichas normas han sido aplicadas en el AMM y los resultados han mostrado que, desde el año 1993, la calidad del aire en el AMM ha inclumpido con las normas de calidad del aire establecidas por la NOM, es decir, año tras año se han registrado días que presentan mediciones de contaminantes mayores a las permitidas.

Es así que este trabajo contribuye con la selección de un modelo de inter-

polación espacial que mejor describa el compartamiento de contaminantes en el AMM. Dichos modelos utilizados en este trabajo, calculan valores de contaminantes en lugares dentro del AMM, donde se desconocen los niveles de contaminación, haciendo uso únicamente de los valores observados en las trece estaciones, luego este trabajo continúa con la creación de mapas que describen el comportamiento de los contaminantes del aire en el AMM, con la premisa de encontrar un modelo que describa lo mejor posible el sistema del aire en el AMM, ya que el organismo encargado de las mediciones de la calidad del aire en el AMM solo cuenta con trece estaciones de monitoreo.

2.1.2 OBJETIVO

El objetivo principal de este trabajo es encontrar un modelo de interpolación espacial que mejor represente los niveles de contaminación del aire en el AMM, para así continuar con la generación de mapas que describan los niveles de contaminación en la ciudad. Además, este trabajo tiene como objetivos secundarios encontrar factores que contribuyan a los niveles altos de contaminación del aire en el AMM, así como la determinación de puntos espaciales críticos dentro de la ciudad donde es importante establecer una nueva estación de monitoreo.

2.1.3 HIPÓTESIS

Es posible estimar niveles de contaminación del aire en puntos geográficos no muestreados dentro del AMM con un nivel de error tolerable, a partir de implementar métodos de interpolación geoespaciales y deterministas, utilizando los datos de las estaciones de monitoreo del SIMA.

2.1.4 ALCANCE

Este trabajo utiliza datos reportados de trece estaciones fijas de monitoreo en el AMM, bajo el cargo del SIMA de la Secretaría de Desarrollo Sustentable del Gobierno de Nuevo León. Algunas estaciones de monitoreo reportan datos desde el año 1993 y actualmente se mantienen operando, aunque algunas se encuentren en estado de mantenimiento.

2.1.5 MOTIVACIÓN

En la Constitución Política de los Estados Unidos Mexicanos se encuentran estipulados el derecho humano a la salud y al medio ambiente saludable, es por esto que es de vital importancia que la normatividad mexicana evolucione para hacer respetar estos derechos. Si bien, el Principio de Progresividad consiste en la obligación del Estado de generar, en cada momento de la historia, una mayor y mejor protección de los derechos humanos, es así que siempre deberá de estar en constante evolución y bajo ninguna justificación en retroceso. Es por esto que se establece una obligación por parte del Estado de monitorear la calidad del aire y de comunicar los resultados a la población (Greenpeace, 2019).

En México el uso de índices de calidad del aire como método de comunicación de riesgo ha evolucionando de manera diferente entre entidades del país, ya que mientras existen entidades como la Ciudad de México y el Estado de México, donde hay antecedentes de intervención que van más allá de los años ochentas y que, actualmente cuentan con una normatividad local al respecto; hay otras entidades como Nuevo León, Baja California, Chiapas, Chihuahua, Coahuila, Durango, Nayarit, Veracruz, Jalisco, Hidalgo, Guanajuato, Queréta-

ro y Oaxaca; donde se han hecho esfuerzos más recientes para desarrollar sus propios índices, pero que carecen de un documento oficial en el que se defina el significado del mismo y los lineamientos para su generación, uso y difusión.

Recientemente, se han encendido alarmas ambientales decretadas en Monterrey, Nuevo León, sobre la mala calidad del aire, que posiciona a la ciudad como una de las más contaminadas del país, en lo que se refiere a material particulado, como el PM_{10} y el $PM_{2.5}$ (Greenpeace, 2019).

En el año 2018, doscientos cuatro días del año superaron los límites máximos de contaminantes en el aire establecidos en las normas ambientales (Organización Mundial de la Salud, 2019). En estos días, contaminantes como $PM_{2.5}$ y PM_{10} ponen en riesgo la salud de la población del AMM, sobre todo si se encuentra expuesta al exterior.

La exposición al exterior durante días que superan el límite máximo de contaminación por periodos cortos o prolongados, puede producir en las personas enfermedades respiratorias y cardiovasculares, como neumonía, asma, gripe, cáncer e incluso la muerte; viéndose afectados principalmente niños y personas mayores (Organización Mundial de la Salud, 2019). Además el estado de Nuevo León es uno de los que reporta mayores promedios de días de hospitalización en pacientes de cero a cuatro años de edad por asma (Greenpeace, 2011).

El impulso de esta tesis está enfocado en el problema de identificar, analizar y proponer un modelo que caracterice a la red de calidad del aire en el AMM, con el fin de entender su comportamiento y evolución. Además de encontrar características importantes en ésta, para así inferir en cuales son las principales fuentes de contaminación del aire en la red.

2.1.6 JUSTIFICACIÓN

Ante la persistente mala calidad del aire que se registra en el SIMA (Secretaría de Desarrollo Sustentable de Nuevo León, 2019), es necesario tomar medidas preventivas contra las causas de la contaminación del aire en el AMM, por lo que se propone como metodología el uso de metodos de interpolación geoespaciales y deterministas, como herramientas para identificar, evaluar y analizar dicha problemática.

Además, la contaminación del aire en las ciudades es un grave problema ambiental, especialmente en las ciudades en desarrollo (Mayer, 1999), como lo es la actual AMM. La ruta de contaminación del aire en la atmósfera urbana consiste en la emisión y transmisión de contaminantes del aire que resultan en la contaminación del ambiente (Alloway y Ayres, 1997). Conocer la calidad del aire en el ambiente permite a las personas residentes del AMM saber si es propicio llevar a cabo actividades en el exterior, esto con el fin de implementar medidas preventivas en días con altas concentraciones de sustancias contaminantes que dañan la salud.

La principal razón que vuelve a este tema relevante es la poca eficiencia de métodos empleados para dar solución a la mala calidad del aire en la zona de AMM.

2.2. Estructura de la Tesis

Este trabajo de tesis está organizado como sigue:

Capítulo 1. Se presenta la introducción de esta tesis, donde se describe la definición

del problema, el objetivo, las hipótesis, alcances, motivación y justificación de la tesis.

Capítulo 2. Se mencionan los antecedentes del marco teórico a seguir en este trabajo y algunas aplicaciones de los métodos usados en este trabajo.

Capítulo 3. Se describe el marco teórico de este trabajo, empezando con el análisis estadístico de los datos haciendo uso de series de tiempo, correlaciones y autocorrelaciones. Luego, se describen los conceptos y definiciones de los métodos que se implementaron en este trabajo.

Capítulo 4. Se describe la metodología implementada a partir del marco teórico antes mencionado y se expone cómo fue la experimentación de este trabajo.

Capítulo 5. Se describen los resultados obtenidos en la experimentación y las conclusiones pertinentes a los resultados, haciendo una comparación con el estado del arte.

CAPÍTULO 3

ANTECEDENTES Y APLICACIONES

En este capítulo se menciona el estado del arte que implementa este trabajo. Primero, se describen los resultados y técnicas que se han aplicado en el AMM para describir los niveles de contaminación del aire y se plantea la razón por la cual en este trabajo se implementan los métodos de interpolación espacial. Luego, se enuncian los antecedentes y algunas aplicaciones en orden cronológico de los métodos de interpolación utilizados en este trabajo: Teselaciones de Voronoi (TV), Funciones de Base Radial (FBR), Distancia Inversa Ponderada (DIP), Kriging Ordinario (KO) y Kriging Universal (KU).

3.1. Calidad del Aire en el AMM

El Sistema Integral de Monitoreo Ambiental (SIMA) del gobierno del estado de Nuevo León, es el organismo gubernamental encargado de reportar y dar a conocer a la población residente del AMM sobre los estándares o niveles de contaminación del aire, que se puede traducir en calidad del aire del AMM.

Actualmente, el SIMA cuenta con trece estaciones de monitoreo de calidad del

aire bajo su responsabilidad, el sistema consiste en recopilar muestras de aire en los monitores y guardarlos en una base de datos. Estas estaciones de monitoreo (no todas) empezaron a trabajar desde 1993, sin embargo por cuestiones técnicas o ambientales los datos no se han podido registrar desde entonces por lo que los datos presenta ausencia de información en el periodo de tiempo 1993–2019.

Una debilidad que presentan las redes de monitoreo (conjunto de estaciones de monitoreo) es que las mediciones tomadas por un medidor solo representa un microambiente geoespacial de contaminantes del aire, es aquí donde entra el problema de SIMA, ya que al contar con trece estaciones no es posible representar la calidad y contaminación del aire sobre toda el AMM.

Se han hecho reportes de la calidad del aire que muestran que el AMM presenta mala calidad del aire desde 1993: y a continuación se mencionan cronológicamente algunos resultados obtenidos sobre ésta. Cabe destacar que aunque se han hecho varias técnicas para describir la calidad del aire en el AMM, hasta la fecha no se ha implementado la idea de interpolar las mediciones de la muestras conocidas en las trece estaciones para tratar de describir la calidad del aire del AMM.

En 1977, se publicó en la revista *Salud Pública de México* en su volumen XIX, el artículo “Información de la calidad en algunas ciudades del país”, donde se menciona que la actividad industrial, la circulación vehicular, la erosión terrestre, el viento, la lluvia, el clima, la topografía y la actividad humana son factores responsables de la calidad del aire.

También, hace referencia a algunas propuestas para estudiar la calidad del aire por parte del gobierno, las cuales se mencionan a continuación:

- A fines de la década 1950 en la Dirección de Higiene Industrial de la Secretaría

de Salubridad y Asistencia, se empezaron diversos estudios sobre la calidad del aire metropolitano en distintas ciudades, incluida Monterrey.

- Más tarde en 1966, se inició un programa de estudio de la contaminación atmosférica que fue reforzado con un convenio entre la Organización Panamericana de la Salud y el Gobierno de los Estados Unidos Mexicanos.
- Hasta el día 31 de enero de 1977 se logró activar el programa en algunos estados de la República Mexicana como son: el Área metropolitana del Valle de México; Guadalajara. Jal.; Monterrey. N.L.; Ciudad Juárez. Chih.; Tijuana y Mexicali, B.C., donde fueron determinados los niveles de concentración de partículas en suspensión. En algunos sitios se investigó la precipitación pluvial, la distribución del tamaño de las partículas, las concentraciones de bióxido de azufre y de ozono.

En resumen, el 31 de enero de 1977 estaban instaladas y en operación cuarenta y ocho unidades muestreadoras de partículas en suspensión en grandes volúmenes de aire, distribuidas en la forma siguiente:

- 22 en el área del Valle de México.
- 10 en Guadalajara. Jal.
- 10 en Monterrey. N.L.
- 2 en Ciudad Juárez. Chih.
- 2 en Tijuana. B.C.
- 2 en Mexicali, B.C.

En la figura 3.1, se presenta un dibujo con las ubicaciones geográficas de diez estaciones de monitoreo del aire en el AMM.

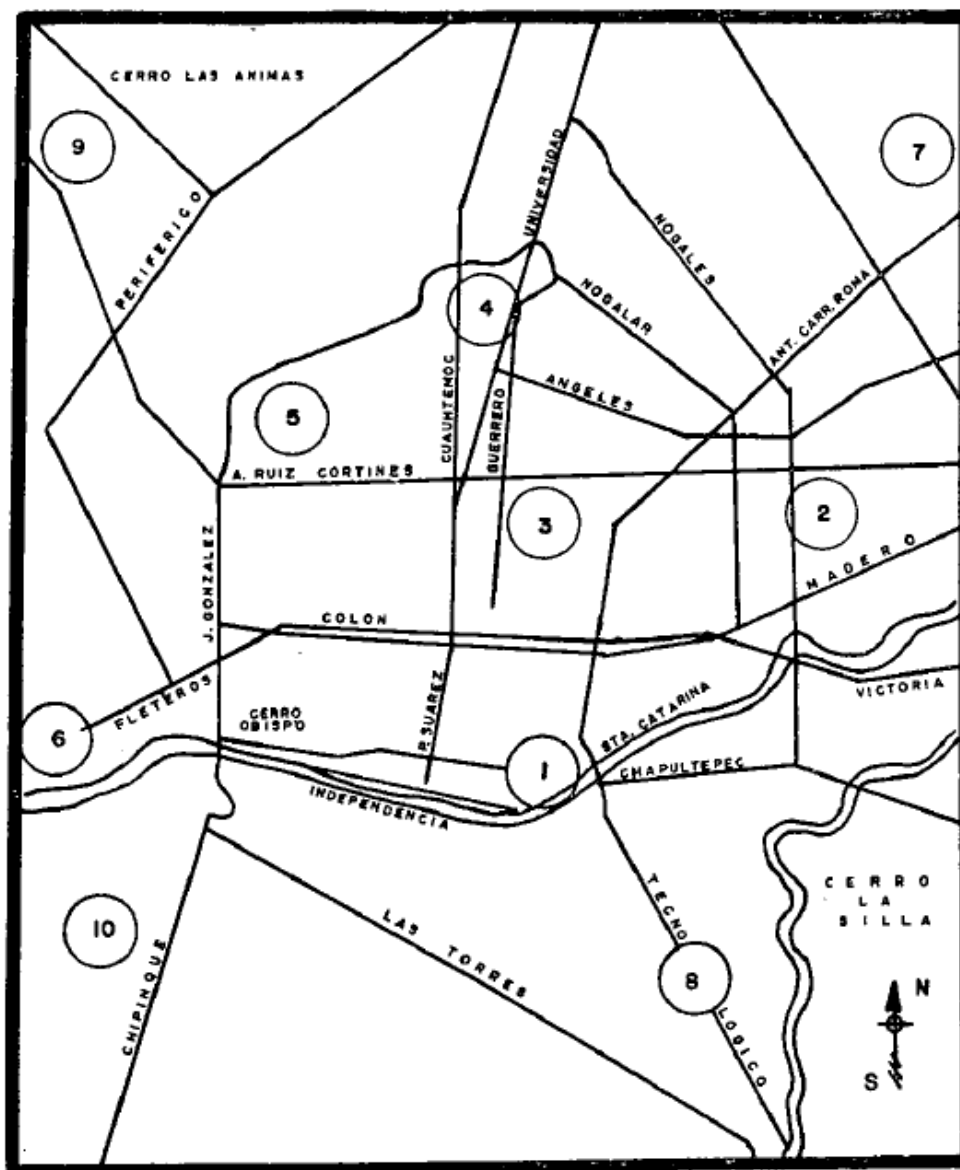


FIGURA 3.1: Red de monitoreo del aire localizado en Monterrey N.L.

- En 2002, como parte de las publicaciones periódicas por la universidad Tecnológico de Monterrey, campus Monterrey, se publicó en la revista *Calidad Ambiental, Elemento Esencial para el Desarrollo Sostenible* en su volumen VII, un análisis que resume diez años de mediciones sobre la calidad del aire del AMM, en el artículo “Análisis de las Mediciones de la Calidad del Aire en el Área Metropolitana de Monterrey: SIMA (1993–2001)”, donde se men-

ciona que las mediciones de la calidad del aire son componentes primordiales para entender, modelar, simular, pronosticar y desarrollar programas para el control de la contaminación atmosférica. También, alude que las mediciones son el registro histórico para verificar y actualizar la efectividad de los diversos programas implantados para la reducción o control de la contaminación urbana.

El programa SIMA del estado de Nuevo León contaba en 2002, con cinco estaciones de monitoreo en: Santa Catarina (suroeste, SO), San Bernabé (noroeste, NO), Obispado (centro, C), San Nicolás (noroeste, NE) y La Pastora (sureste, SE). Los contaminantes monitoreados en las cinco estaciones son: Monóxido de carbono (CO), partículas menores a diez micras (PM_{10}), ozono (O_3), óxidos de nitrógeno (NO y NO_2) y bióxido de azufre (SO_2); y los parámetros meteorológicos medidos son: dirección del viento, velocidad del viento y temperatura ambiental.

Para informar en forma clara y sencilla los niveles de contaminación existentes se utilizó el *Índice Metropolitano de Calidad del Aire* (IMECA). Un valor de cien puntos IMECA se establece cuando el contaminante se encuentra en una concentración igual a la norma de calidad. Véase tabla 2.1 a tabla 2.7.

TABLA 3.1: Cumplimiento a las Normas de Calidad de Aire en la AMM (1993–2000)

	Max/Min	Norma de Calidad			Incumplimientos	
		ppmv	Horas	Número	Máximo ppmv	Región del máximo
CO	1.17	11.0	8	6	12.9	NE
PM_{10}	1.8	150.00 (a)	24	336	271 (a)	SO
O_3	1.78	0.11	1	152	0.194	NO
NO_2	1.34	0.21	1	1	0.281	C
SO_2	0.68	0.13	24	0	0.089	NE
(a) La norma y las concentraciones medidas para PM_{10} están en ug/m^3 .						

En la tabla 3.1 los contaminantes de PM_{10} y Ozono, son los contaminantes que

han tenido los máximos puntos IMECA. Esto es, en situaciones ambientales severas en Monterrey los IMECAs se aproximan a 200 puntos.

TABLA 3.2: Días Arriba de la Norma en la AMM (1993–2000)

Año	Totales *	PM ₁₀	Ozono	Regiones con más días
1993	80	44	37	SO/NO
1994	99	84	21	SO/NO
1995	31	27	4	NO/SO
1996	62	40	24	NO/SO
1997	38	8	32	SO/NO
1998	35	21	14	SO/NO
1999	93	84	8	NE/NO
2000	30	24	8	SO/NO
2001	113	91	12	SO/NO
Total	581	423	160	SO/NO
Promedio	65	47	18	

En la tabla 3.2 se reportaron el número de días que las PM₁₀ y Ozono han estado arriba de la norma en Monterrey desde 1993 hasta 2001. Además, indica las regiones con más días arriba de la norma, Ozono ha estado arriba de la norma dieciocho días al año en promedio, mientras que PM₁₀ ha rebasado la norma en cuarenta y siete días por año. En total, algunos de los contaminantes ha rebasado la norma en sesenta y cinco días por año, es decir, estos datos implican que la contaminación atmosférica en la AMM es alta 18 % de los días del año (65 días), ya sea debido a Ozono o PM₁₀.

TABLA 3.3: Cumplimiento Anual de IMECAs Máximos en la AMM (1993–2000)

	Promedio	Máximo de IMECAs Máximos			Mínimo de IMECAs Máximos		
		Valor	Región	Año	Valor	Región	Año
CO	60	120	NE	1996	30	SE	2000
PM ₁₀	152	255	C	2001	75	SE	1997
O ₃	117	192	NO	2001	72	NE	1998
NO ₂	69	117	C	1993	31	NO	1999
SO ₂	32	69	NE	1994	10	C	2000

TABLA 3.4: Regiones con Mayores IMECAs Máximos en la AMM (1993–2000)

Año	CO	PM ₁₀	Ozono	NO ₂	SO ₂
1993	NE/C	SO/NO	SO/C	C/SO	NE/C
1994	NE/SO	NO/SO	SO/C	C/SO	NE/C
1995	NE/C	NO/NE	SO/NO	SO/NO	NE/C
1996	NE/C	NO/NE	SE/SO	C/NE	NE/NO
1997	NE/C	NO/C	SE/SO	NE/C	NE/NO
1998	NE/SO	NO/SO	NO/SO	C/NO	NE/C
1999	NE/SO	NE/C	SO/NO	NE/SO	NE/NO
2000	NE/SO	SO/NO	NO/SO	NE/SO	NE/NO
2001	NE/C	C/SE	NO/SO	NE/SO	NE/NO

La tabla 3.3 muestra los máximos y mínimos que ocurrieron de 1993 a 2001 reportando la región y el año de ocurrencia del evento. La tabla 3.4 muestra las dos regiones con mayores IMECAs máximos por año y para cada contaminante.

TABLA 3.5: Cumplimiento Mensual de la Calidad del Aire en la AMM (1993–2000)

	Relación	Máximo de Promedios			Mínimo de Promedios		
	Max/Min	ppmv	Mes	Regiones	ppmv	Mes	Regiones
CO	3.6	1.800	Enero	C/NE	0.5	Julio	NE/NO
PM ₁₀	2.6	92 (a)	Enero	NO/SO	35 (a)	Septiembre	C/SE
O ₃	2.7	0.032	Abril	SE/SO/NO	0.012	Enero	C/NE
NO ₂	5.4	0.038	Enero	C/SO	0.007	Julio	SE/NO
SO ₂	5.0	0.025	Julio	NE	0.005	Junio	SE
(a) La norma y las concentraciones medidas para PM ₁₀ están en ug/m^3 .							

TABLA 3.6: Comportamiento Hora a Hora de la Calidad del Aire en la AMM (1993–2000)

	Relación	Máximo de Promedios			Mínimo de Promedios		
	Max/Min	ppmv	Hora	Regiones	ppmv	Hora	Regiones
CO	5.0	2.0	7 – 9	C/NE	0.4	15 – 17	NO/SE
PM ₁₀	5.4	150 (a)	9 – 10	NO/SO	28 (a)	17 – 18	SE/C
O ₃	5.4	0.054	12 – 13	NE/SO	0.010	5 – 7	C/SO
NO ₂	5.2	0.042	8 – 10	C/SE	0.008	15 – 16	SE/NO
SO ₂	9.3	0.028	11 – 13	NE	0.003	6 – 16	SE
(a) La norma y las concentraciones medidas para PM ₁₀ están en ug/m^3 .							

La tabla 3.5 muestra los máximos y mínimos de los promedios mensuales de cada contaminante de 1993 a 2000, señalando el mes y región donde sucedió el máximo y mínimo. En la tabla 3.6 se muestran las concentraciones promedio hora a hora del periodo que se monitorearon los contaminantes primarios tiene un comportamiento basado en los niveles a tiempo real de: emisiones de aire, condiciones meteorológicas y reactividad química del contaminante para el AMM.

En resumen de las tablas anteriores se puede concluir que:

1. Tabla 3.1: Todos los contaminantes han rebasado la norma excepto el bióxido de azufre; los contaminantes con los máximos más elevados son las PM_{10} y Ozono, los máximos de Ozono y PM_{10} han llegado hasta 80 % arriba de la norma, y solamente la región SE ha cumplido con las normas.
2. Tabla 3.2: La contaminación atmosférica en Monterrey está arriba de la norma de calidad del aire en un 18 % de los días del año debida a Ozono o PM_{10} . En promedio la norma de Ozono ha sido rebasada la norma diecinueve días por año y las partículas cuarenta y siete días por año.
3. Tablas 3.3 y 3.4: En promedio el Ozono es alrededor de 120 puntos y para PM_{10} es 150. También, altos niveles de Ozono se favorecen en las estaciones SO y NO. Altos niveles de PM_{10} suceden en las estaciones NO y SO. Altos niveles de CO y SO_2 suceden en las estaciones NO y C. El Ozono y las PM_{10} se han incrementado fuertemente en el 2001.
4. Tabla 3.5: Las altas concentraciones de Monóxido de carbono, PM_{10} y Bióxido de nitrógeno sucede en diciembre–enero y las bajas en julio–agosto. Mientras que el Ozono presenta altas concentraciones en abril–mayo y sus bajas en diciembre–enero.
5. Tabla 3.6: Las altas concentraciones de Monóxido de carbono, PM_{10} y Bióxido de nitrógeno, suceden al iniciar la mañana y después del atardecer, y las bajas concentraciones en las últimas horas antes de anocheecer. Mientras tanto, el Ozono tiene sus altas concentraciones al iniciar la tarde y sus bajas concentraciones desde el atardecer y durante la noche. En promedio, la contaminación atmosférica en la AMM es al menos cinco veces mayor al iniciar la mañana con respecto al atardecer.

En 2017, el Proyecto de Norma Oficial Mexicana PROY-NOM-172-SEMARNAT-2017, donde se establecen los lineamientos para la obtención y comunicación

del *Índice de Calidad del Aire y Riesgos a la Salud*, fue aprobado por el Comité Consultivo Nacional de Normalización de Medio Ambiente y Recursos Naturales. Dos años más tarde, después de haberse cumplido el procedimiento establecido en la Ley Federal sobre Meteorología y Normalización para la elaboración de normas oficiales mexicanas, el Comité Consultivo Nacional de Normalización de Medio Ambiente y Recursos Naturales aprobó la Norma Oficial Mexicana NOM-172-SEMARNAT-2019, derivada del proyecto PROY-NOM-172-SEMARNAT-2017 el 10 de octubre de 2019.

Con la implementación de los lineamientos de la Norma Oficial Mexicana NOM-172-SEMARNAT-2019 para la obtención y comunicación del *Índice de Calidad del Aire y Riesgos a la Salud*, se establece un método único de cálculo y los lineamientos de difusión que deberán aplicar los gobiernos estatales responsables del monitoreo de la calidad del aire, ya que la importancia del *Índice de Calidad del Aire y Riesgos a la Salud* no sólo consiste en informar a la población sobre el estado de la calidad del aire (buena, aceptable, mala, muy mala y extremadamente mala), sino también sobre el nivel de riesgo asociado (probables daños a la salud, dependiendo si el riesgo es bajo, moderado, alto, muy alto o extremadamente alto) y las recomendaciones de las acciones a adoptar (medidas para reducir la exposición), es decir, se busca que la información que reciba la población no solamente se refiera a la calidad del aire, sino que le permita actuar con oportunidad para proteger su salud.

Actualmente, SIMA cuenta con trece estaciones de monitoreo, de las cuales se puede observar su ubicación geográfica en la figura 3.2. Además en la tabla 3.7 se muestran las características de la red de monitoreo de calidad del aire por estación, se puede ver su ubicación geográfica en coordenadas, su elevación con respecto al nivel del mar, el municipio en el que se encuentra y la fecha desde cuando entró en operación cada estación.

TABLA 3.7: Características de las estaciones de monitoreo del SIMA

No.	Municipio	Zona	Simbología	Coordenadas Geográficas	Elevación	Operando desde
1	Guadalupe, Nuevo León	Sureste	SE	25.668 N, 100.249 W	492 msnm	20/Nov/1992
2	San Nicolás de los Garza, Nuevo León	Noreste	NE	25.75 N, 100.255 W	476 msnm	20/Nov/1992
3	Monterrey, Nuevo León. Obispado	Centro	CE	25.67 N, 100.338 W	560 msnm	20/Nov/1992
4	Monterrey, Nuevo León. Sn. Bernabé	Noroeste	NO	25.757 N, 100.366 W	571 msnm	20/Nov/1992
5	Santa Catarina, Nuevo León	Suroeste	SO	25.676 N, 100.464 W	694 msnm	20/Nov/1992
6	García, Nuevo León	Noroeste 2	NO2	25.783 N, 100.586 W	716 msnm	24/Jul/2009
7	Escobedo, Nuevo León	Norte	N	25.800 N, 100.344 W	528 msnm	22/Dic/2009
8	Apodaca, Nuevo León	Noreste 2	NE2	25.777 N, 100.188 W	432 msnm	01/Jun/2011
9	Juárez, Nuevo León	Sureste 2	SE2	25.646 N, 100.096 W	387 msnm	01/Oct/2012
10	San Pedro Garza García, Nuevo León	Suroeste 2	SO2	25.665 N, 100.413 W	636 msnm	06/Feb/2014
11	Cadereyta de Jiménez, Nuevo León	Sureste 3	SE3	25.36 N, 99.9955 W	340 msnm	21/Ago/2017
12	UANL, San Nicolás, Nuevo León	Norte 2	N2	25.5749 N, 100.2489 W	630 msnm	04/Oct/2017
13	Monterrey, Nuevo León	Sur	S	25.7295 N, 100.3099 W	520 msnm	04/Oct/2017

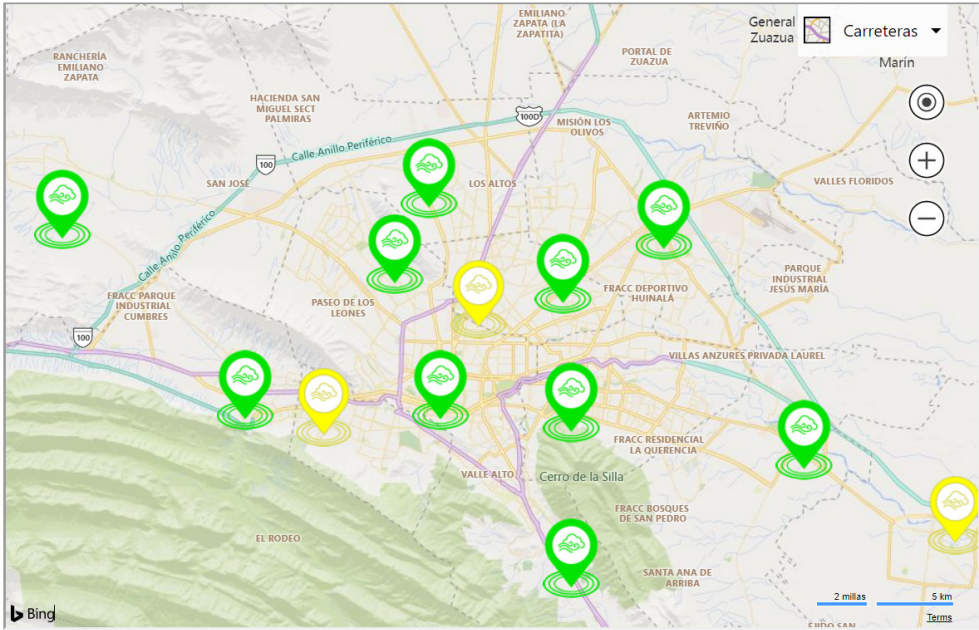


FIGURA 3.2: Mapa del AMM con las trece estaciones meteorológicas marcadas por un símbolo de posición geográfica (Secretaría de Desarrollo Sustentable de Nuevo León, 2020)

TABLA 3.8: Variables y contaminantes medidos por SIMA

Nombre del contaminante	Nomenclatura SIMA	Nomenclatura MEEAMM	Unidad de medición
Partículas menores a 10 micras	PM ₁₀	PM ₁₀	Ugr /m ³
Partículas menores a 2.5 micras	PM _{2.5}	PM _{2.5}	Ugr /m ³
Ozono	O ₃	O ₃	ppb
Óxido Nítrico	NO	NO	ppb
Bióxido de Nitrógeno	NO ₂	NO ₂	ppb
Oxidos de Nitrógeno	NO _x	NO _x	ppb
Bióxido de Azufre	SO ₂	SO ₂	ppb
Monóxido de Carbono	CO	CO	ppb
Temperatura ambiental	TOUT	Temperature	°C
Precipitación pluvial	RAINP	Rainfall	mm/h
Humedad relativa	HR	Humidity	%
Presión barométrica	PRES	Pressure	mmHg
Radiación solar	SR	Solar	kWh/m ²
Velocidad del viento	WSR	Velocity	Km/h
Dirección del viento	WDR	Direction	Grado azimutal

Abreviación del título de esta tesis Modelos Estadístico Espacial de contaminantes del aire en el AMM (MEEAMM)

En las trece estaciones con las que cuenta el SIMA, se miden y registran quince

variables, las cuales se muestran en la figura 3.8; en la cual se muestra el nombre de la variable, la nomenclatura asignada por SIMA, la nomenclatura asignada para este trabajo de tesis y que se emplea en el código fuente (Serna, 2020) (MEAMM, pos sus siglas de Modelo Espacial de contaminantes del Área metropolitana de Monterrey) y la unidad en la que se mide cada variable.

Partiendo de las concentraciones base indicadas en la tabla 3.9 se crea el Índice AIRE y SALUD con base en la Norma Oficial Mexicana PROY-NOM-172-SEMARNAT-2017 (ver tabla 3.9).

TABLA 3.9: Concentraciones base para el cálculo del *Índice AIRE y SALUD* para cada contaminante

Contaminante	Concentración base
PM ₁₀	Concentración promedio móvil ponderado de 12 horas
PM _{2.5}	Concentración promedio móvil de 12 horas
O ₃	Concentración promedio móvil de 8 horas
CO	Concentración promedio móvil de 8 horas
NO ₂	Concentración promedio horaria
SO ₂	Concentración promedio de 24 horas

Las bandas de calidad del aire y riesgo que componen el *Índice AIRE y SALUD* se construyen considerando los intervalos de concentración señalados en la Norma Oficial Mexicana PROY-NOM-172-SEMARNAT-2017, según aplique al contaminante criterio. Particularmente, los límites superiores del intervalo de la banda “Aceptable” concuerdan con los valores establecidos en las Norma Oficial Mexicana NOM-020-SSA1-2014.

TABLA 3.10: Obtención del *Índice AIRE y SALUD* para PM₁₀

Calidad del aire	Nivel de riesgo	Intervalo de PM ₁₀
Buena	Bajo	≤ 50
Aceptable	Moderado	$> 50, \leq 75$
Mala	Alto	$> 75, \leq 155$
Muy Mala	Muy Alto	$> 155, \leq 235$
Extremadamente Mala	Extremadamente Alto	> 235

TABLA 3.11: Obtención del *Índice AIRE y SALUD* para PM_{2.5}

Calidad del aire	Nivel de riesgo	Intervalo de PM _{2.5}
Buena	Bajo	≤ 25
Aceptable	Moderado	$> 25, \leq 45$
Mala	Alto	$> 45, \leq 79$
Muy Mala	Muy Alto	$> 79, \leq 147$
Extremadamente Mala	Extremadamente Alto	> 147

TABLA 3.12: Obtención del *Índice AIRE y SALUD* para O₃

Calidad del aire	Nivel de riesgo	Intervalo de O ₃
Buena	Bajo	≤ 0.051
Aceptable	Moderado	$> 0.051, \leq 0.070$
Mala	Alto	$> 0.070, \leq 0.092$
Muy Mala	Muy Alto	$> 0.092, \leq 0.114$
Extremadamente Mala	Extremadamente Alto	> 0.114

TABLA 3.13: Obtención del *Índice AIRE y SALUD* para NO₂

Calidad del aire	Nivel de riesgo	Intervalo de NO ₂
Buena	Bajo	≤ 0.107
Aceptable	Moderado	$> 0.107, \leq 0.210$
Mala	Alto	$> 0.210, \leq 0.230$
Muy Mala	Muy Alto	$> 0.230, \leq 0.250$
Extremadamente Mala	Extremadamente Alto	> 0.250

TABLA 3.14: Obtención del *Índice AIRE y SALUD* para SO₂

Calidad del aire	Nivel de riesgo	Intervalo de SO ₂
Buena	Bajo	≤ 0.008
Aceptable	Moderado	$> 0.008, \leq 0.110$
Mala	Alto	$> 0.110, \leq 0.165$
Muy Mala	Muy Alto	$> 0.165, \leq 0.220$
Extremadamente Mala	Extremadamente Alto	> 0.220

TABLA 3.15: Obtención del *Índice AIRE y SALUD* para CO

Calidad del aire	Nivel de riesgo	Intervalo de CO
Buena	Bajo	≤ 8.75
Aceptable	Moderado	$> 8.75, \leq 11.00$
Mala	Alto	$> 11.00, \leq 13.30$
Muy Mala	Muy Alto	$> 13.30, \leq 15.50$
Extremadamente Mala	Extremadamente Alto	> 15.50

La difusión de riesgos relacionada al *Índice AIRE y SALUD* consistirá en el establecimiento de cinco bandas que estarán asociadas a cinco colores (verde, amarillo, naranja, rojo y morado) como se describe en la tabla 3.16. El Índice AIRE Y SALUD sólo tiene fines de información para prevenir a la población en

una ciudad o localidad en una hora determinada, estos riesgos estan asociados a los niveles de contaminación de cada contaminante descritos en la tablas 3.10, 3.11, 3.12, 3.13, 3.14, y 3.15.

TABLA 3.16: Categorías del *Índice AIRE y SALUD*

Calidad del aire	Nivel de riesgo	Color
Buena	Bajo	Verde
Aceptable	Moderado	Amarillo
Mala	Alto	Naranja
Muy Mala	Muy Alto	Rojo
Extremadamente Mala	Extremadamente Alto	Morado

3.2. Métodos de Interpolación

Esta sección presenta brevemente los diferentes métodos de interpolación utilizados en este trabajo de tesis, como son los ya conocidos; Teselaciones de Voronoi, Funciones de Base Radial, Distancia Inversa Ponderada, Kriging Universal y Kriging Ordinario. Cabe destacar que los métodos de interpolación espacial mencionados difieren en sus supuestos, perspectiva local o global y naturaleza determinista o estocástica.

3.2.1 TESELACIONES DE VORONOI

Las Teselaciones de Voronoi o Diagramas de Voronoi (TV) son estructuras geométricas que aparecen con regularidad en la naturaleza, por esta razón se han redescubierto varias veces a lo largo de la historia, reciben su nombre en honor al matemático ruso Georgy Voronoi, también son conocidas como Polígonos

de Thiessen (por el meteorólogo estadounidense Alfred Thiessen), Teselación de Dirichlet (por el matemático alemán Peter Gustav Lejeune Dirichlet), Celdas de Wigner-Seitz (por el físico matemático húngaro Eugene Wigner y el físico estadounidense Frederick Seitz) o Zonas de Brillouin (por el físico francés León Nicolás Brillouin (Aurenhammer y Klein, 2000)).

En 1664, el filósofo y matemático francés Descartes afirmó que el sistema solar se compone de *vórtices*; considerados como cualquier tipo de flujo circular o rotatorio que posee vorticidad. Las ilustraciones de Descartes muestran una descomposición del espacio en regiones convexas, donde cada una de éstas está compuesta de la materia que gira alrededor de una de las estrellas fija (Aurenhammer y Klein, 2000).

En 1850, Dirichlet formalizó el estudio de formas cuadráticas definidas positivas (forma especial de un Diagrama de Voronoi) en \mathbb{R}^2 y \mathbb{R}^3 (Dirichlet, 1850). Cuatro años después en 1854, el médico inglés John Snow (padre de la epidemiología moderna), hizo un análisis del brote de cólera que azotó la ciudad de Londres, mediante el uso del método geográfico (Doval, 2003).

En 1908, Voronoi precisó el estudio de formas cuadráticas definidas positivas en \mathbb{R}^n . Un año más tarde en 1909, el geólogo, cristalógrafo y mineralogista ucraniano Boldyrev (Anatoly Kapitonovich) estimó las reservas de minerales en un depósito usando la información obtenida de taladros, nombrando a las regiones de Voronoi *áreas de polígonos de influencia*.

En 1911, Thiessen analizó datos meteorológicos de estaciones pluviométricas, donde a las regiones las nombró *polígonos de Thiessen*. En 1927, el cristalógrafo suizo Paul Niggli desarrolló la teoría en Cristalografía (Manzoor *et al.*, 2014). Dos años después en 1929, el matemático ruso Boris Deloné (Delaunay) fue el primero en acuñar el término *Dominio de Dirichlet* y el término *Región de Voronoi*.

En 1933, Wigner y Seitz describieron las regiones de Voronoi para los puntos dispuestos en una retícula en el espacio tridimensional (Senechal, 1993). En 1949, el demógrafo estadounidense Donald Bogue, usó los polígonos de Voronoi definidos alrededor de centros metropolitanos.

En 1965, G. S. Brown estimó la intensidad de población de los árboles en un bosque, definiendo una región de Voronoi para un árbol individual, llamándola el *área potencialmente disponible* (APA) de un árbol (Brown, 1965). En 1985, L. Hoofd et al. definieron las regiones de Voronoi con respecto a los centros de los capilares en secciones de tejido (Turek *et al.*, 1992). Dos años después en 1987, V. Icke y R. Van de Weygaert realizaron una partición del espacio por medio de un proceso de diagramas de Voronoi (Icke *et al.*, 1987).

3.2.2 FUNCIONES DE BASE RADIAL

Las Funciones de Base Radial (FBR) comprenden un gran grupo de interpoladores exactos, los cuales utilizan una *ecuación básica* que depende de la distancia entre los puntos de muestreo y el punto interpolado. Dicha ecuación toma en cuenta que la función de interpolación minimiza una función que representa alguna medida de suavidad (Aguilar *et al.*, 2005).

Un tipo particular de las funciones de base radial son las de *soporte global*, las cuales son funciones infinitamente diferenciables y con valor real no nulo en todos los puntos de su dominio, y donde algunas contienen un parámetro libre, llamado *parámetro de forma*. El uso de este tipo de funciones básicas produce una matriz de interpolación que sirve para interpolar suavemente y generar zonas continuas sobre superficies discontinuas (Arias *et al.*, 2009).

Algunos tipos de funciones de base radial de soporte global más populares en

la literatura que se usan en este trabajo son: *Multicuadrática* (FBRM), *Inversa multicuadrática* (FBRIM), *Gaussiana* (FBRG), *Biarmónica* o *Lineal* (FBRL), *Triarmónica* o *Cúbica* (FBRC) y *Thin Plate Spline* (FBRTPS). Todas ellas han tenido buena aceptación debido a que el sistema asociado de ecuaciones lineales resulta ser invertible, incluso si la distribución de los puntos no presenta regularidad (Arias *et al.*, 2009).

En 1977, A. Talmi y G. Gilat plantearon los antecedentes matemáticos para la construcción de funciones de base radial en su trabajo *Method for Smooth Approximation of Data* (Talmi y Gilat, 1977). La aplicación de estos métodos de interpolación a grandes volúmenes de datos presenta una serie de problemas, como el alto costo computacional, un aumento en la inestabilidad de la solución encontrada y la influencia de todos los puntos de muestreo en cada valor interpolado (Lazzaro y Montefusco, 2002).

En 1982, R. Franke recomendó que el BRFM es el que proporciona los mejores resultados en términos de evaluación estadística y visual de la superficie modelizada (Franke, 1982). Ya que una ecuación de topografía basada en la suma multicuadrática aplica un concepto geométrico simple, donde la suavidad y la forma de la transición entre los puntos de datos están controladas principalmente por las características del cuadrícula básico utilizado en la suma (Hardy, 1971).

En 1990, R. L. Hardy comentó que el BRFM funciona bien como método de interpolación, cuando las mediciones de potencial o temperatura en la superficie de la tierra se obtienen en estaciones meteorológicas dispersas (Hardy, 1990).

En 1992, F. Girosi mencionó que las funciones de base radial son un enfoque adecuado cuando la aproximación a las llamadas técnicas de aprendizaje de las redes neuronales conducen a problemas de interpolación de muy alta dimensión con datos dispersos, debido a su disponibilidad en dimensiones arbitrarias y a

su suavidad (Giroi, 1992).

En 1996, T. Sonar utilizó funciones de base radial para la reconstrucción local de soluciones dentro de algoritmos que resuelven leyes de conservación numéricamente hiperbólicas. Hasta entonces era habitual emplear una aproximación polinómica de bajo orden (en su mayoría lineal) para este propósito; pero resultó que las funciones de base radial, especialmente las FBRTPS, ayudan a mejorar la precisión de los métodos de volumen finito, sobre todo para resolver las ecuaciones hiperbólicas, debido a su capacidad para aproximarse localmente con gran precisión (Sonar, 1996).

En 2002, A. Kremper, T. Schanze y R. Eckhorn usaron datos provenientes de una pantalla que muestra la lectura de una cámara que sirve como ojo de un robot. Las imágenes representan objetos que el robot debería reconocer, por ejemplo, un muro con el que no debería estrellarse, otro robot, un humano o cualquier otro objeto. Sin embargo, se tuvo el problema que cada uno de estos datos debía interpolarse para que el robot fuera capaz de reconocer otras situaciones similares, ya que las imágenes tenían variaciones como: ángulos, escalas, etc. Para este problema se encontró que usando interpolaciones de FBR, el robot lograba identificar una mayor cantidad de datos (Kremper *et al.*, 2002).

3.2.3 DISTANCIA INVERSA PONDERADA

La interpolación de Distancia Inversa Ponderada (DIP) se aplica ampliamente por los científicos de la tierra (Ware *et al.*, 1991) y en la cartografía automatizada, ya que realiza el cálculo de una superficie cuadrículada bidimensional a partir de un conjunto de puntos dispersos con valores conocidos (Armstrong y Marciano, 1994). En las geociencias, la idea de usar DIP se encuentra en-

caminada a evaluar la contaminación en un punto de referencia, a partir de las mediciones realizadas por las estaciones de monitoreo (Vogl *et al.*, 2012). DIP se basa en la premisa de que las predicciones son una combinación lineal de datos disponibles (Xie *et al.*, 2011), donde los pesos dependen de la distancia entre las estaciones de monitoreo con un exponente arbitrario y en consecuencia, son exógenos con respecto a los datos (Armstrong y Marciano, 1994).

En 1996, P. Bartier y C. Keller presentaron una metodología de interpolación que aborda explícitamente la variación de la superficie a través de límites poligonales temáticos. El método de interpolación de DIP se implementó para permitir a los usuarios definir el grado esperado de brusquedad de la superficie a lo largo de los límites temáticos utilizando una matriz de transición. El procedimiento se demostró utilizando como caso de estudio, una superficie interpolada de geoquímica de níquel modificada por límites estratigráficos de terreno (Bartier y Keller, 1996).

En 2004, R. L. Hough realizó una evaluación de riesgo de exposición a metales en subgrupos de población que viven y cultivan alimentos en sitios urbanos, donde se utilizó la interpolación por DIP en las concentraciones de metales en las verduras, para evaluar el riesgo de exposición a las poblaciones humanas (Hough *et al.*, 2004).

En 2005, J. Currie y M. Neidel examinaron el impacto de la contaminación del aire en la muerte infantil en California durante la década de 1990. Aquí utilizaron DIP en niveles de contaminación de CO del aire y descubrieron que las reducciones en el monóxido de carbono durante la década de 1990 salvaron aproximadamente 1000 vidas infantiles en California (Currie y Neidell, 2005).

En 2009, O. Babak y C. V Deutsch presentaron un enfoque para integrar controles estadísticos, como la mínima varianza de error en la interpolación

de DIP. Mencionando que el exponente óptimo y el número de datos pueden calcularse global o localmente, y que las medidas de incertidumbre y suavidad local pueden derivarse de estimaciones de DIP (Babak y Deutsch, 2009).

En 2012, F. Chen y C. Liu utilizaron DIP para estimar la distribución de lluvia en Taiwán, con un total de cuarenta y seis estaciones y datos de lluvia entre 1981 y 2010, de las cuales doce estaciones pertenecientes a la Asociación de Riego de Taichung (TIA), se usaron para la validación cruzada. Encontrando que los parámetros óptimos para DIP en la interpolación de datos de lluvia tienen un radio de influencia de hasta 10–30 km en la mayoría de los casos y que, usando DIP en estos datos de interpolación pueden obtenerse resultados más precisos durante la estación seca que en la temporada de inundaciones. En conclusión, obtuvieron que los altos valores de coeficiente de correlación (más de 0.95) confirmaron a DIP como un método adecuado de interpolación espacial para predecir los datos probables de lluvia en Taiwán (Chen y Liu, 2012).

3.2.4 KRIGING ORDINARIO

El término *Krigeage* fue acuñado por P. Carlier en reconocimiento de D.G. Krige por su innovación pionera para estimar las concentraciones de oro y otros metales en el suelo (Oliver y Webster, 2015).

El método Kriging Ordinario (KO) predice valores en sitios no visitados a partir de datos de muestra dispersos, basados en un modelo estocástico de variación espacial continua; teniendo en cuenta conocimiento de la variación espacial representada en el *variograma* o *función de covarianza*.

En 1963, G. Matheron en su libro *Principles of geostatistics* introdujo al idioma

inglés el termino *Kriging* (Matheron, 1963) y dos años después, en 1965, en su tesis doctoral colocó la técnica dentro del marco de la teoría de procesos aleatorios (Matheron, 1965). El trabajo de Matheron no estaba aislado, ya que trabajos como los de; Andréi Kolmogorov (1939, 1941), Herman Wold (1938) y Wiener (1964), ya habían llegado cerca del kriging, pero en el tiempo más que en el espacio (Cressie, 1990).

En 1990, M. Oliver y R. Webster comentaron que los sistemas de información geográfica pueden mejorarse agregando procedimientos para el análisis espacial geoestadístico a las estaciones de monitoreo existentes, ya que los datos distribuidos espacialmente se comportan más como variables aleatorias, donde la teoría de variables regionalizadas proporciona un conjunto de métodos estocásticos que minimizan el error de predicción (Oliver y Webster, 1990).

En 1994, R. Dimitrakopoulos y X. Luo aludieron que se puede usar una representación general de ‘anisotropía zonal’ para una covarianza espacio-temporal positiva. Se demostró que una covarianza espacio-temporal definida estrictamente positiva, se puede asegurar la unicidad de la solución del sistema Kriging (Dimitrakopoulos y Luo, 1994). Aquí resultados de Dimitrakopoulos y Luo se basaron en el enfoque de Buxton y Pate (1994), así como en los de Bilonick (1985, 1988).

En 2004, L. Bel utilizó un estimador no paramétrico y un modelo de transporte químico para producir mapas de concentración de Ozono sobre el área de París y compararlos con los mapas obtenidos con los métodos de Kriging clásicos. Bel mostró que al tener pocas estaciones KO resulta ser malo interpolando cuando se aplica con los estimadores clásicos (Bel, 2004).

En 2007, M. Villatoro, C. Henríquez y F. Sancho hicieron una comparación entre KO y DIP de mediciones de materiales en el suelo. Para el KO se calcularon los semivariogramas y también se determinó que el modelo esférico fue

el de mejor ajuste. Ambos métodos tuvieron un desempeño similar, el KO fue superior al predecirlo de una mejor manera (Villatoro *et al.*, 2007).

En 2015, O. Delgado y J. Martínez emplearon el método KO para construir el mapa de ruido de la Ciudad de Cuenca, Ecuador; su trabajo partió con la determinación de los sitios de muestreo sobre la base de la densidad de tráfico, luego se registraron las mediciones de ruido ambiente y posteriormente se sistematizó y evaluó la información levantada a través del método KO, con lo cual se elaboró el mapa de ruido de la ciudad (Delgado y Martínez, 2015).

Formas más elaboradas de Kriging se encuentran desarrollando para abordar problemas cada vez más complejos en la ingeniería petrolera, minería y geología, meteorología, ciencia del suelo, agricultura de precisión, contaminación, salud pública, pesca, ecología vegetal y animal, teledetección e hidrología.

3.2.5 KRIGING UNIVERSAL

El método Kriging Universal (KU) a diferencia del método KO supone que la media de los datos $m(x)$ tiene una dependencia funcional de la ubicación espacial y puede ser aproximada por un modelo con la ecuación (Kumar, 2007):

$$\mu(x) = \sum_{i=1}^k a_i f_i(x), \quad (3.1)$$

donde a_i son los coeficientes a estimar a partir de los datos, f_i es una función básica de coordenadas espaciales y k es el número de funciones utilizadas en el modelo.

En 1988, S. Dingman, D. Seely y R. Reynolds estudiaron las mediciones de los equilibrios químicos y del agua para las cuencas de drenaje y los cuerpos de agua superficiales. Mencionando que los errores en sus predicciones usando

KU pueden deberse a fallas de los medidores individuales para recolectar la cantidad de precipitación que realmente cae, errores del operador o fallas de la red de regado para muestrear adecuadamente la región de interés. Sus resultados mostraron que el KU a diferencia de KO presenta un error considerablemente más paqueño, por lo que es preferible como base para estimaciones puntuales (Dingman *et al.*, 1988).

En 2002, P. Kanaroglou, N. Soulakellis y N. Sifakis demostraron que el KU puede usarse para obtener estimaciones razonables para datos con valores faltantes para la evaluación de la distribución espacial de contaminantes del aire. Usaron imágenes de satélite implementando el método de Análisis de Textura Diferencial, el cual se utiliza para medir el espesor óptico del aerosol en lo visible, que se correlaciona altamente con la calidad del aire. Con este método, la presencia de nubes o cambios en la cobertura del suelo producen parches de valores perdidos, los que se recuperaron usando KU (Kanaroglou *et al.*, 2002).

En 2011, L. Mercer et al. mencionaron que el método KU es un buen método de interpolación para, valga la redundancia, interpolar mediciones de nitrógeno (NO_x) en Los Ángeles, C.A. (Mercer *et al.*, 2011). Dos años después en 2013, P. Sampson et. al. presentaron un modelo de KU para interpolar datos de partículas menores a dos punto cinco micras ($\text{PM}_{2.5}$), de estaciones de monitoreo en los EE. UU., reflejando un nivel muy alto de precisión de predicción bajo el método de validación cruzada (Sampson *et al.*, 2013).

En 2016, I. Kiš et al. hicieron una comparación entre los metodos KO Y KU, sus resultados arrojaron que el método KU predice mejor que KO cuando los datos presentan una tendencia espacial lineal, ya que esta técnica se desarrolló precisamente para identificar y calcular tendencias en los datos, es decir, para describir la regresión y dar mejores resultados de mapeo (Kiš, 2016).

CAPÍTULO 4

MARCO TEÓRICO

En este capítulo se exponen los conceptos que explican el comportamiento de los datos y la suma importancia de conocer las características y comportamiento para hacer uso de los datos en la experimentación de este trabajo de tesis. Primero se habla sobre las *series de tiempo* y las características que éstas deben cumplir para poder decir que estadísticamente los datos con los que contamos son útiles para la experimentación. Luego, se exponen los conceptos de *autocorrelación*, *autocovarianzas* y *variograma*, y cómo estos conceptos ayudan a entender el comportamiento de los datos y a mejorar los resultados de la experimentación.

4.1. Análisis Estadístico

Se utilizan herramientas estadísticas para conocer el comportamiento de los datos, primero se usan series de tiempo para graficar los datos de contaminantes por estaciones y observar las características de éstas. Luego, se verifica si las series cumplen con la propiedad de ser estacionarias en el tiempo. Después se calculan las correlaciones y autocorrelaciones de los contaminantes por estación.

y por último, se calcula el variograma, el cual se utiliza como parámetro para encontrar el mejor método de interpolación de los métodos geoespaciales.

4.1.1 SERIES DE TIEMPO

Una *serie de tiempo* es una secuencia cronológica de observaciones sobre una variable de interés (Montgomery *et al.*, 2015). Por ejemplo, la figura 4.1 muestra el rendimiento del mercado de los valores del tesoro de EE. UU. con vencimiento constante a 10 años desde abril de 1953 hasta diciembre de 2006. Este gráfico se llama *diagrama de serie de tiempo*.

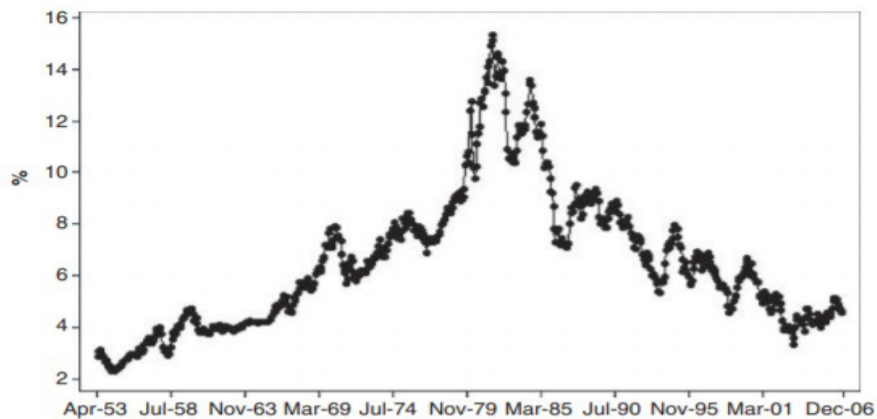


FIGURA 4.1: Gráfico de la serie de tiempo del rendimiento del mercado de valores del tesoro de EE. UU. con vencimiento constante a 10 años (Montgomery *et al.*, 2015)

Las mediciones de la variable de interés a estudiar se pueden recopilar en períodos de tiempo igualmente espaciados, como es típico en la mayoría de las series de tiempo. Los tiempos de recopilación pueden ser diarios, semanales, mensuales, trimestrales o anuales; pero se puede utilizar cualquier intervalo de tiempo, por ejemplo, en este trabajo las mediciones son hora a hora. Además,

los datos pueden ser *instantáneos*, como el nivel de un contaminante químico del aire en el momento en que se mide; *acumulativo*, como las ventas totales de un producto durante un mes; o puede ser una *estadística*, que de alguna manera refleja la actividad de la variable durante el período de tiempo.

La razón por la cual el registro de mediciones de una variable es tan importante se da porque la predicción de eventos futuros de dicha variable es un excelente aporte en muchos tipos de procesos de planificación y toma de decisiones, con aplicación en áreas como (Montgomery *et al.*, 2015):

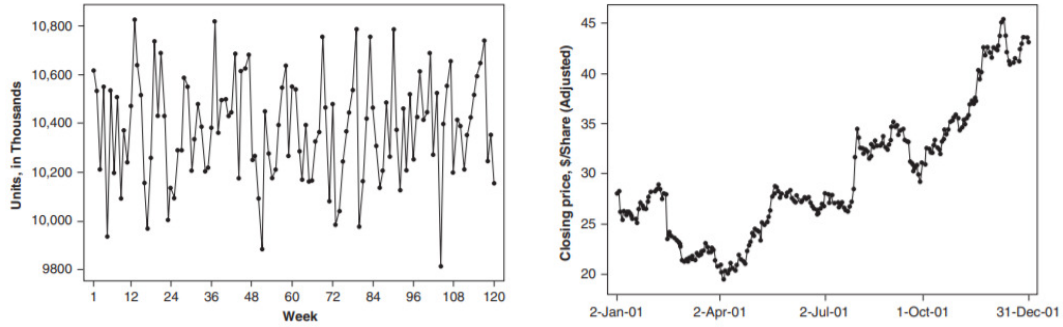
- Márketing
- Finanzas y gestión de riesgos
- Ciencias económicas
- Control de procesos industriales
- Demografía

4.1.2 ESTACIONARIEDAD

Un tipo importante de series de tiempo es una serie de tiempo *estacionaria*. Se dice que una serie de tiempo es *estrictamente estacionaria* si sus propiedades no se ven afectadas por un cambio en el origen del tiempo. Es decir, si la distribución de probabilidad conjunta de las observaciones $y_t, y_{t+1}, \dots, y_{t+n}$ es exactamente la misma que la distribución de probabilidad conjunta de las observaciones $y_{t+k}, y_{t+k+1}, \dots, y_{t+k+n}$, entonces la serie temporal es estrictamente estacionaria.

La figura 4.2 (a) es un ejemplo de serie de tiempo estacionaria, ya que parece variar entorno a un nivel fijo, ésta es una característica de las series de tiempo

estacionarias. Por otro lado, la figura 4.2 (b) tiende a deambular a la deriva sin un nivel fijo obvio, este es un comportamiento típico de una serie de tiempo no estacionaria.



(a) Serie estacionaria (Montgomery *et al.*, 2015), (b) Serie no estacionaria (Montgomery *et al.*, 2015)

FIGURA 4.2: Series de tiempo

Cuando $n = 0$, el supuesto de estacionariedad significa que la distribución de probabilidad de y_t es la misma para todos los períodos de tiempo y puede escribirse como $f(y)$.

Estacionariedad implica un tipo de equilibrio estadístico o estabilidad en los datos (Montgomery *et al.*, 2015), es decir, la serie temporal tiene una media constante definida de la manera habitual como:

$$\mu_y = E[y] = \int_{-\infty}^{\infty} y f(y) dy, \quad (4.1)$$

y la varianza constante definida como:

$$\sigma_y^2 = \text{Var}[y] = \int_{-\infty}^{\infty} (y - \mu_y)^2 f(y) dy. \quad (4.2)$$

La media muestral y la varianza muestral se utilizan para estimar estos parámetros. Si las observaciones en la serie de tiempo son y_1, y_2, \dots, y_t , entonces la

media muestral es:

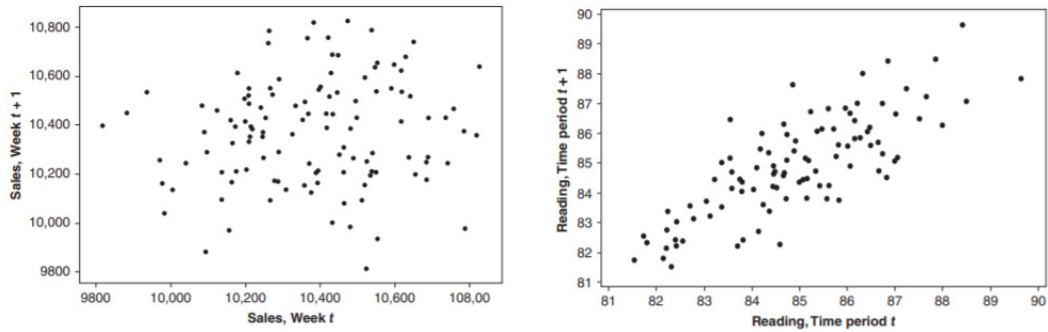
$$\bar{y} = \hat{\mu}_y = \frac{1}{t} \sum_{t=1}^t y_t, \quad (4.3)$$

y la varianza muestral es:

$$s^2 = \hat{\sigma}_y^2 = \frac{1}{t} \sum_{t=1}^t (y_t - \bar{y})^2. \quad (4.4)$$

4.1.3 AUTOCOVARIANZAS Y AUTOCORRELACIONES

Si una serie de tiempo es estacionaria, esto significa que la distribución de probabilidad conjunta de cualquiera de las dos observaciones, por ejemplo, y_t y y_{t+k} , es la misma para dos períodos de tiempo t_y y t_{t+k} que están separados por el mismo intervalo de tamaño k . El intervalo k se llama *retraso*.



(a) Datos no correlacionados (Montgomery *et al.*, 2015) (b) Datos correlacionados (Montgomery *et al.*, 2015)

FIGURA 4.3: Datos correlacionados y no correlacionados

Las figuras 4.3 (a) y 4.3 (b) son diagramas de dispersión con un retraso $k = 1$. Ambos diagramas de dispersión se construyeron trazando y_{t+1} versus y_t . La figura 4.3 (a) exhibe poca estructura, ya que los pares trazados de observaciones adyacentes y_t, y_{t+1} parecen no estar *correlacionados*, es decir, el valor de y en el

período actual no proporciona ninguna información útil sobre el valor de y que se observará en el próximo período. Por otro lado, en figura 4.3 (b), los pares de observaciones adyacentes y_{t+1}, y_t están positivamente correlacionados, es decir, un valor pequeño de y tiende a ser seguido en el siguiente período por otro valor pequeño de y , y un valor grande de y tiende a ser seguido inmediatamente por otro valor grande de y .

La covarianza entre y_t y su valor en otro período de tiempo, digamos, y_{t+k} se llama *autocovarianza* en el retraso k , definido por:

$$\gamma_k = \text{Cov}(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)]. \quad (4.5)$$

La colección de los valores de γ_k , $k = 0, 1, 2, \dots$ se llama *Función de autocovarianza*, donde la *autocovarianza* en el retraso $k = 0$ es la *varianza* de la serie de tiempo, es decir, $\gamma_0 = \sigma_y^2$, que es constante para una serie de tiempo estacionaria. Además, el coeficiente de autocorrelación en el intervalo k para una serie temporal estacionaria se calcula como:

$$\rho_k = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E[(y_t - \mu)]^2 E[(y_{t+k} - \mu)^2]}} = \frac{\text{Cov}(y_t, y_{t+k})}{\text{Var}(y_t)} = \frac{\gamma_k}{\gamma_0}. \quad (4.6)$$

La colección de los valores de ρ_k , $k = 0, 1, 2, \dots$ se llama *Función de autocorrelación (ACF)*, donde $\rho_0 = 1$. Además, el ACF es independiente de la escala de medición de la serie de tiempo. También, $\rho_k = \rho_{-k}$, es decir, el ACF es simétrico alrededor de cero, por lo que solo es necesario calcular la mitad positiva (o negativa).

Si una serie de tiempo tiene función de autocovarianza y media finita, se dice que es estacionaria de segundo orden (o débilmente estacionaria de orden dos). Si además, la distribución de probabilidad conjunta de las observaciones en todo momento es multivariante normal, entonces eso es suficiente para dar como resultado una serie de tiempo estrictamente estacionaria.

Es necesario estimar la autocovarianza y los ACF a partir de un momento de longitud finita, digamos, y_1, y_2, \dots, y_t . La estimación habitual de la función de autocovarianza se calcula como:

$$c_k = \hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y}), \quad k = 0, 1, \dots, k, \quad (4.7)$$

y el ACF es estimado por la función de autocorrelación de muestra (o muestra de ACF) calculada como:

$$r_k = \hat{\rho}_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, k. \quad (4.8)$$

4.1.4 VARIOGRAMA

A menudo en la práctica no hay una demarcación clara entre un proceso estacionario y uno no estacionario para muchas series de tiempo del mundo real. Por ello, una herramienta de diagnóstico adicional que es muy útil es el *Variograma*.

Suponga que las observaciones de la serie de tiempo están representadas por y_t . El variograma G_k mide las variaciones de las diferencias entre las observaciones que están separadas a una distancia k , en relación con la variación de las diferencias que están separadas por una unidad de tiempo. El variograma se define como:

$$G_k = \frac{\text{Var}(y_{t+k} - y_t)}{\text{Var}(y_{t+1} - y_t)}, \quad k = 1, 2, \dots \quad (4.9)$$

y los valores de G_k se representan en función del retraso k . Si la serie temporal es estacionaria, resulta que:

$$G_k = \frac{1 - \rho_k}{1 - \rho_1}, \quad (4.10)$$

para una serie temporal estacionaria $\rho_k \rightarrow 0$ a medida que k aumenta, por lo que cuando el variograma se traza contra el retraso k , G_k alcanzará una asíntota $\frac{1}{(1-\rho_1)}$. Sin embargo, si la serie temporal no es estacionaria, G_k aumentará monótonamente.

4.2. Métodos de Interpolación

Sea $P_n = \{p_1, p_2, \dots, p_n\}$, $n \geq 2$ el conjunto de puntos muestrales (estaciones de monitoreo) en el espacio \mathbb{R}^2 o *plano Euclideo*, etiquetados mediante las coordenadas $(x_1, y_1), \dots, (x_n, y_n)$. Los n puntos son distintos en el sentido de que $(x_i, y_i) \neq (x_j, y_j) \forall p_i, p_j \in P_n$.

Definición 4.1. (Distancia Euclidiana). Sean p y p^* dos puntos arbitrarios en el plano Euclidiano con coordenadas (x, y) y (x^*, y^*) respectivamente. Entonces, la *distancia Euclidiana* entre p y p^* está dada por:

$$d(p, p^*) = \|p - p^*\| = \sqrt{(x - x^*)^2 + (y - y^*)^2}. \quad (4.11)$$

Nota: Se considera a partir de ahora que se trabaja en el espacio Euclidiano con la *métrica Euclidiana*.

4.2.1 DETERMINISTAS

La interpolación mediante métodos de interpolación deterministas indica los valores de un punto desconocido a través de una combinación ponderada linealmente de un conjunto de puntos conocidos (puntos de muestra). Los métodos de interpolación deterministas suponen que la superficie que se interpola debe ser la de una variable dependiente de la ubicación. En esta sección se definen los

métodos de interpolación deterministas que se utilizan en este trabajo y son: Diagramas de Voronoi, Funciones de Base Radial y el método de Distancia Inversa Ponderada.

4.2.1.1 DIAGRAMAS DE VORONOI

Definición 4.2. (Diagrama de Voronoi (DV)). Sea el conjunto $P_V = \{p_1, p_2, \dots, p_n\} \subset \mathbb{R}^2$, con $2 < n$ y $x_i \neq x_j \forall i, j$. Llamemos a la región dada por:

$$V(p_i) = \{x : \|x - x_i\| \leq \|x - x_j\|, i \neq j\} \forall x_i, x_j \in P_V, \quad (4.12)$$

como el *Polígono de Voronoi* asociado a p_i (o el Polígono de Voronoi de p_i y al conjunto dado por

$$V = \{V(p_1), \dots, V(p_n)\} \quad (4.13)$$

el *Diagrama de Voronoi* generado por P_V (o el diagrama de Voronoi de P_V).

En la definición 4.2, se observa de la ecuación 4.12 que está definida en términos de desigualdad (\leq), por lo que un polígono de Voronoi es un *conjunto cerrado*. Dado que un polígono de Voronoi, éste contiene su frontera ($\partial V(p_i)$) y podemos establecer las siguientes definiciones.

Definición 4.3. (Arista de Voronoi). Puesto que el diagrama de Voronoi contiene a su frontera $\partial V(p_i)$, ésta puede consistir en segmentos de recta, semirectas o rectas infinitas, las cuales llamamos *Aristas de Voronoi*. Denotamos a la arista de los polígonos de Voronoi p_i y p_j como $a(p_i, p_j)$.

Definición 4.4. (Vértice de Voronoi). Llamamos *Vértice de Voronoi* a un punto extremo de una arista de Voronoi o se puede definir también como un punto compartido por tres o más polígonos de Voronoi. Denotamos a el vértice de Voronoi del polígono de Voronoi p_i como v_i .

Definición 4.5. (Red de Voronoi). El diagrama de Voronoi puede ser definido por la unión de las aristas de Voronoi, es decir, $\bigcup_{i=1}^n \partial V(p_i)$ en lugar del conjunto P_V . Llamamos a red (*mallá*), conformada por la unión de aristas de Voronoi como la *Red de Voronoi*.

4.2.1.2 FUNCIONES DE BASE RADIAL

Definición 4.6. (Funciones de Base Radial (FBR)). Sea el conjunto $P_{\mathcal{F}} = \{p_1, p_2, \dots, p_n\} \subset \mathbb{R}^2$, con $x_i \neq x_j \forall i, j$. Llamemos *Función de Base Radial* a una función perteneciente a una gran familia de interpoladores exactos que usan una *ecuación básica* que es una función que depende de la distancia entre el punto interpolado y los puntos de muestreo $P_{\mathcal{F}}$. Los valores de predicción por una FBR se pueden expresar como la suma de dos componentes:

$$\hat{Z}(x) = \sum_{i=1}^n a_i f_i(x) + \sum_{j=1}^n b_j \psi(d_j), \quad (4.14)$$

donde $\psi(d_j)$ es la ecuación básica que depende de la distancia d_j , donde d_j es la distancia desde los puntos de muestra $P_{\mathcal{F}}$ a el punto interpolado x y $f_i(x)$ es un polinomio de grado $< m$, con m el número de puntos a interpolar. Los coeficientes a_i y b_j se calculan mediante la resolución del sistema de $n + m$ ecuaciones lineales, con n el número de puntos conocidos $P_{\mathcal{F}}$ utilizados en la interpolación de la siguiente manera:

$$\begin{aligned} \hat{Z}(x_k) &= \sum_{i=1}^m a_i f_i(x_k) + \sum_{j=1}^n b_j \psi(d_{jk}), \quad k = 1, 2, \dots, n \\ \sum_{j=1}^n b_j f_k(x_j) &= 0, \quad k = 1, 2, \dots, m. \end{aligned} \quad (4.15)$$

En este trabajo de tesis se utilizan las siguientes clases de BRF para la interpolación de datos dispersos n -dimensionales a un dominio m -dimensional; Multicuadrática (M), Inversa multicuadrática (IM), Gaussiana (G), Lineal (L),

Cúbica (C), Quintica (Q) y Thin-plate splines (TPS). La función de base radial utilizada para cada caso da las siguientes expresiones:

- M: $\psi(d) = \sqrt{d^2 + c^2}$
- IM: $\psi(d) = \sqrt{d^2 + c^2}^{-1}$
- G: $\psi(d) = e^{-\sqrt{d^2 + c^2}}$
- L: $\psi(d) = d$
- C: $\psi(d) = d^3$
- Q: $\psi(d) = d^5$
- TPS: $\psi(d) = d^2 + c^2 \ln(cd)$,

donde d es la distancia de los puntos muestra $P_{\mathcal{F}}$ a la ubicación de predicción x y c es un factor de suavizado.

4.2.1.3 DISTANCIA INVERSA PONDERADA

Definición 4.7. (Distancia Inversa Ponderada (DIP)). Sea el conjunto $P_{\mathcal{D}} = \{p_1, p_2, \dots, p_n\} \subset \mathbb{R}^2$, con $x_i \neq x_j \forall i, j$. Llamemos función de *Distancia Inversa Ponderada* a una función de la forma:

$$\hat{Z}(x) = \frac{\sum_{i=1}^n w_i z_i}{\sum_{i=1}^n w_i} \quad (4.16)$$

$$w_i = d_i^{-u},$$

donde $Z(x)$ es el valor predicho en el punto x , Z_i es el valor del punto muestra p_i , d_i es la distancia entre el punto p_i y el punto x y w_i es el peso asignado al punto p_i .

Los valores de mayor ponderación se asignan a valores más cercanos al punto interpolado. A medida que la distancia aumenta, el peso disminuye (Shepard, 1968), y u es el poder de ponderación que decide cómo disminuye el peso a medida que aumenta la distancia.

4.2.2 GEOESTADÍSTICOS

La interpolación mediante métodos de interpolación geoestadísticos o probabilísticos se resumen en el método de interpolación *Kriging*. La palabra Kriging (expresión anglosajona) proviene del nombre del geólogo sudafricano Danie G. Krige, cuyo trabajo aportó a la predicción de reservas de oro, realizados en la década de los años cincuentas y éste suele considerarse como el pionero en los métodos de interpolación geoespacial.

Kriging encierra un conjunto de métodos de predicción espacial que se centran en la minimización del error cuadrático medio de predicción. Al igual que con DIP, el estimador de Kriging viene dado por una combinación lineal de los valores observados y los pesos. Dependiendo de las propiedades estocásticas de los campos aleatorios, se aplican diferentes tipos de Kriging. En esta sección se definen los métodos de interpolación geoestadísticos que se utilizan en este trabajo y son: Kriging Ordinario y Kriging Universal.

4.2.2.1 KRIGING ORDINARIO

Definición 4.8. (Kriging Ordinario (KO)). Sea el conjunto $P_O = \{p_1, p_2, \dots, p_n\} \subset \mathbb{R}^2$, con $x_i \neq x_j \forall i, j$. Supongase que se hacen mediciones de la variable de interés Z en los puntos $p_i \in P_O \forall i$, es decir, se tienen $Z(x_1), \dots, Z(x_n)$, y se desea predecir el punto $Z(x^*)$, en el punto $x^* \notin P_O$ donde no hubo medición. Llamemos *Kriging*

Ordinario al método que predice con una combinación lineal las n variables aleatorias como:

$$\hat{Z}(x^*) = \lambda_1 Z(x_1) + \lambda_2 Z(x_2) + \dots + \lambda_n Z(x_n) = \sum_{i=1}^n \lambda_i Z(x_i), \quad (4.17)$$

donde λ_i representan los pesos de los valores de los puntos muestrales y x_i son los valores observados en los puntos p_i .

Dichos pesos se calculan en función de la distancia entre los puntos muestrales (P_O) y el punto predicho x^* . La suma de los pesos debe ser igual a uno para que la esperanza del predictor sea igual a la esperanza de la variable. Esto último se conoce como *propiedad de insesgamiento*.

Estadísticamente la propiedad de insesgamiento es:

$$\begin{aligned} E \left[\sum_{i=1}^n \lambda_i Z(x_i) \right] &= m, \\ \sum_{i=1}^n \lambda_i E[Z(x_i)] &= m, \\ \sum_{i=1}^n \lambda_i m &= m, \\ m \sum_{i=1}^n \lambda_i &= m \Rightarrow \sum_{i=1}^n \lambda_i = 1. \end{aligned}$$

Se dice que $\hat{Z}(x^*)$ es el mejor predictor lineal en este caso, ya que los pesos se obtienen de tal manera que minimicen la varianza del error de predicción, es decir, que minimicen la expresión dada por:

$$V[\hat{Z}(x^*) - Z(x^*)]. \quad (4.18)$$

Esta característica es distintiva de los métodos Kriging, ya que métodos de interpolación deterministas no garantizan varianza mínima de predicción (Samper y Carrera, 1990). La estimación de los pesos se obtiene resolviendo:

$$\text{Min. } V[\hat{Z}(x^*) - Z(x^*)]$$

$$\text{s. a. } \sum_{i=1}^n \lambda_i = 1.$$

Los pesos de KO se derivan de la *función de semivariancia*. Los parámetros de la función de semivariancia y el *efecto de pepita* se pueden estimar mediante una *función de semivariación empírica* (Webster y Oliver, 2007). Un estimador imparcial de la función de semivariancia es la mitad de la diferencia cuadrática promedio entre los valores de datos dado por:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2, \quad (4.19)$$

donde $\gamma(h)$ es el valor de semivariancia en el intervalo de distancia h , $N(h)$ es el número de pares de muestras dentro del intervalo de distancia h y $z(x_i + h)$ y $z(x_i)$ son valores de muestra en dos puntos separados por el intervalo de distancia h .

4.2.2.2 KRIGING UNIVERSAL

En los supuestos de KO se ha asumido que la variable es estacionaria, pero en la mayoría de los casos la variable no satisface esta condición y se caracteriza por exhibir una tendencia. Para tratar este tipo de variables es frecuente descomponer la variable $Z(x)$ en la suma de la tendencia, tratada como una función determinística más una componente estocástica estacionaria de media cero.

Definición 4.9. (Kriging Universal (KU)). Sea el conjunto $P_U = \{p_1, p_2, \dots, p_n\} \subset \mathbb{R}^2$, con $x_i \neq x_j \forall i, j$. Llamemos *Kriging Universal* al método que predice como la suma de la tendencia, tratada a manera de una función determinística más una componente estocástica estacionaria de media cero las n variables aleatorias, por ejemplo:

$$\hat{Z}(x) = m(x) + \epsilon(x), \quad (4.20)$$

con $E[\epsilon(x)] = 0$, $V[\epsilon(x)] = \sigma^2$ y $E[Z(x)] = m(x)$.

La tendencia puede expresarse mediante:

$$m(x) = \sum_{l=1}^p a_l f_l(x),$$

donde las funciones $f_l(x)$ son conocidas y p es el número de términos empleados para ajustar $m(x)$. Por lo tanto, el predictor de KU se define como:

$$\hat{Z}(x^*) = \sum_{i=1}^n \lambda_i Z(x_i). \quad (4.21)$$

CAPÍTULO 5

METODOLOGÍA Y EXPERIMENTACIÓN

Este capítulo comienza abordando los temas explicados en el marco teórico y su implementación a los datos en este trabajo de tesis. Primero, se muestra en análisis estadístico de los datos, número de variables, cantidad de registros, cantidad de datos faltantes y cantidad de datos atípicos o anómalos; acompañado de descripciones visuales como series de tiempo por estación de cada variable, seguido de matrices de correlación entre estaciones por variable. Después, se menciona cómo se eliminan datos no deseables y se completan dichos registros para así, por último, tener una base de datos decente para llevar a cabo la experimentación y obtención de resultados discutidos en el capítulo siguiente.

5.1. Análisis Estadístico y Limpieza de la Base de Datos

Las estaciones de monitoreo indican las tendencias puntuales en la calidad del aire, pero sólo representan microambientes alrededor de ellas, por lo que se utilizan *métodos de interpolación espacial* para calcular los niveles de contami-

nantes en los lugares donde no hay estaciones de monitoreo.

La interpolación es el proceso de predecir valores en sitios sin muestrear. Los métodos de interpolación espacial incorporan información sobre la posición geográfica de los puntos de muestra, de modo que los puntos más cercanos entre sí tienen una mayor correlación y similitud que los que están más lejos. En este trabajo se evalúan los siguientes métodos de interpolación: Diagramas de Voronoi (DV), Distancia Inversa Ponderada (DIP), Funciones de Base Radial (FBR), Kriging Ordinario (KO) y Kriging Universal (KU).

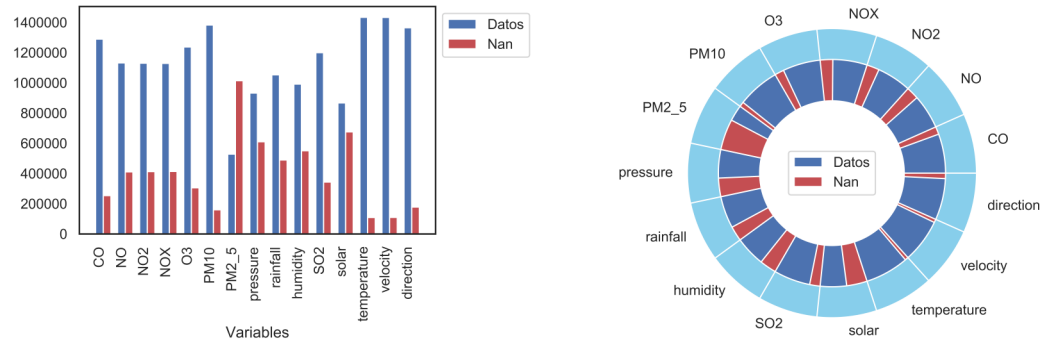
La recolección de los datos se hace a partir del promedio de mediciones por hora. La base de datos cuenta con lo siguiente:

- Trece estaciones de monitoreo.
- 1,542,696 mediciones del promedio por hora (desde 1993-01-01 00:00:00 hrs. hasta 2018-12-31 23:00:00 hrs.).
- Mediciones de quince variables por estación.

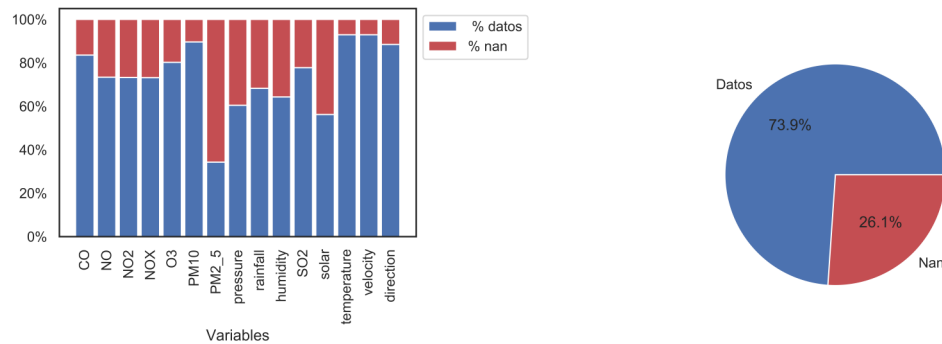
Luego se cuenta el número de datos faltantes en la base de datos por variable.

TABLA 5.1: Contaminantes y variables que recopilan las estaciones de monitoreo del SIMA con los nombres correspondientes de las variables en el código fuente (Serna, 2020)

Variable	Nomenclatura MEEAMM	Cantidad de NaN
Tiempo	timestamp	0
Monóxido de Carbono	CO	253,474
Óxido Nítrico	NO	410,737
Bióxido de Nitrógeno	NO ₂	412,339
Óxidos de Nitrógeno	NO _x	413,766
Ozono	O ₃	305185
Partículas menores a 10 micras	PM ₁₀	159,985
Partículas menores a 2.5 micras	PM _{2.5}	1,014,253
Presión Barométrica	pressure	610,362
Precipitación Pluvial	rainfall	489,331
Humedad Relativa	humidity	550,968
Bióxido de Azufre	SO ₂	343,180
Radiación Solar	solar	675,701
Temperatura Ambiental	temperature	109,143
Velocidad del Viento	velocity	109,590
Dirección del Viento	direction	178,126
Total		7,554,588



(a) Cantidad de datos conocidos y datos desconocidos por variable (b) Porcentaje de datos conocidos y datos desconocidos por variable



(c) Porcentaje de datos conocidos y datos desconocidos por variable (d) Porcentaje de datos conocidos y datos desconocidos

FIGURA 5.1: Datos conocidos y desconocidos (NaN) registrados por SIMA

De las quince variables con las que cuenta SIMA, que son: CO, NO, NO₂, NO_x, O₃, PM₁₀, PM_{2.5}, *Presión atmosférica*, *Precipitación pluvial*, *Humedad relativa*, SO₂, *Radiación solar*, *Temperatura ambiental*, *Velocidad del viento* y *Dirección del viento*; se analizan cada una por individual. Lo primero que se hace es seleccionar los datos desde el año 2016, ya que algunas estaciones fueron puestas a trabajar hasta el año 2017; para recopilar la información de las variables en estas estaciones se hace el promedio de las estaciones más cercanas sin valores registrados para, posteriormente asignárselos.

Después de identificar la cantidad de datos faltantes, se prosigue con los valo-

res que no son faltantes pero son no permitidos por estar fuera del rango de medición, esto es, el valor existe, aunque el valor de la medición fue provocado por una falla en el sensor de la estación de monitoreo. Los valores mínimos y máximos para cada variable se muestran en la tabla 5.2.

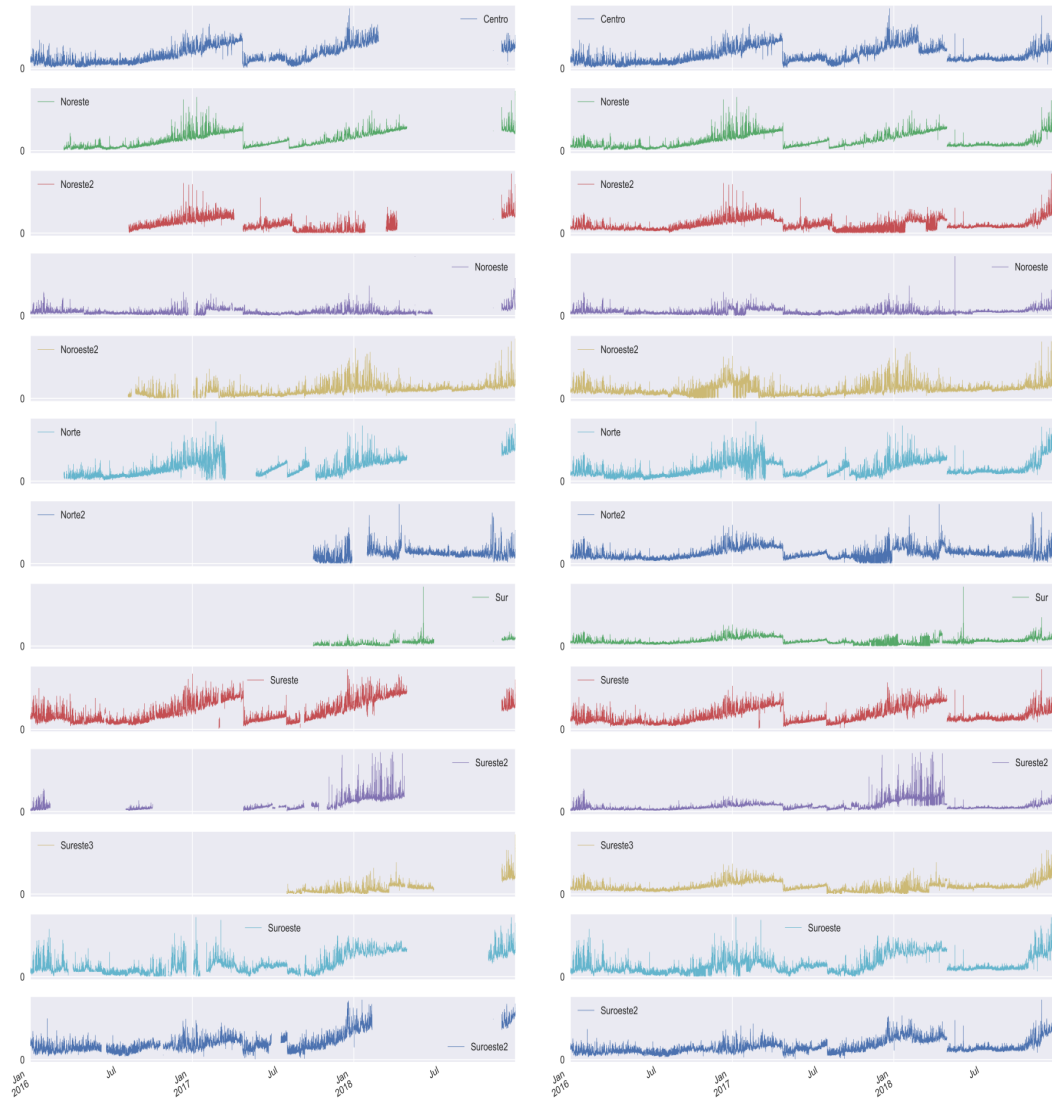
TABLA 5.2: Valores fuera de rango por falla en analizadores y sensores de los diferentes parámetros en las estaciones de monitoreo atmosférico

Parámetro	Valores no permitidos	
	Menor a la lectura	Mayor a la lectura
Partículas menores a 10 micras	2	850
Partículas menores a 2.5 micras	2	850
Ozono	1	200
Óxido Nítrico	1	350
Bióxido de Nitrógeno	1	150
Oxídos de Nitrógeno	1	350
Bióxido de Azufre	1	200
Monóxido de Carbono	0.05	15
Temperatura ambiental	-15	50
Precipitación pluvial	0	—
Humedad relativa	0	100
Presión barométrica	650	750
Radiación solar	0	1.2
Velocidad del viento	0	60
Dirección del viento	0	360

Después de identificar los valores faltantes y los valores no permitidos, se eliminan de la base de datos. Para completar los registros eliminados lo que se hace es interpolar los datos mediante la función `interpolate` con el método `time` cuando el valor fue un valor no permitido, ya que la función `interpolate` consiste en seleccionar los valores al tiempo $t - 1$ y al tiempo $t + 1$ para asignar

un valor en el registro al tiempo t mediante una interpolación lineal. Para el caso de las estaciones en las que no se cuenta con información al tiempo $t - 1$ y $t + 1$, pues no se encontraban en operación la estación, lo que se hace es calcular el promedio por variable de las estaciones más cercanas a la estación que aún no está en operación, asignándole el valor promedio calculado hora a hora.

5.1.1 MONÓXIDO DE CARBONO (CO)



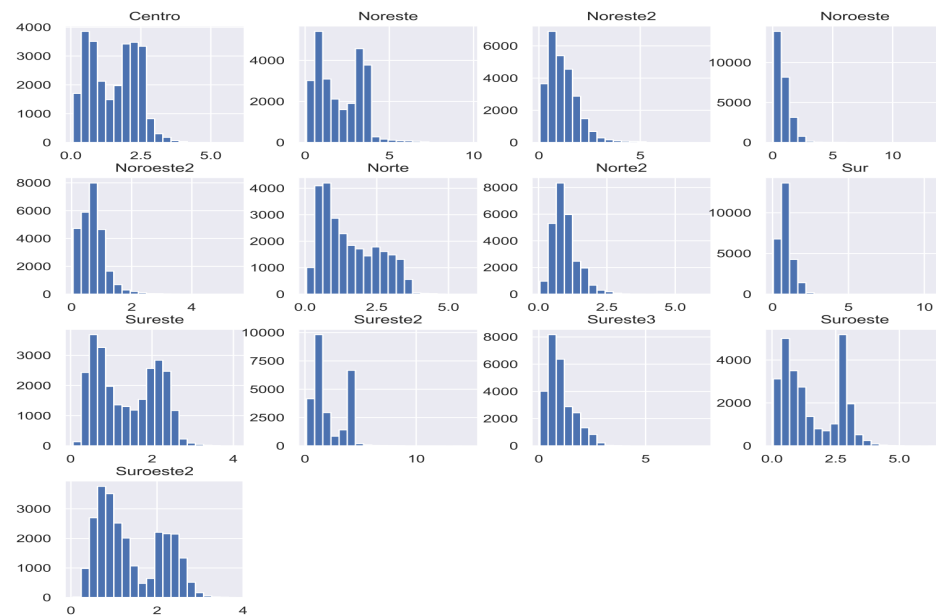
(a) Serie de tiempo de CO original

(b) Serie de tiempo de CO rellenando datos desconocidos

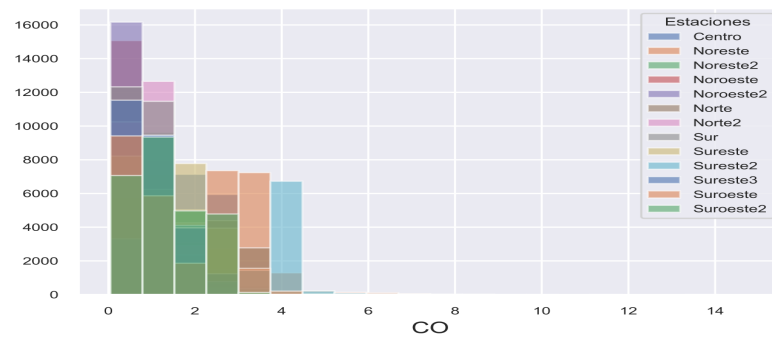
FIGURA 5.2: Series de tiempo de CO desde 2016 hasta 2018

La variable CO presenta datos faltantes durante intervalos de tiempo en los tres años seleccionados (ver figura 5.2 (a)), tales valores se completan haciendo el promedio de las estaciones cercanas a la estación con los datos faltantes y

se asigna el valor promedio a ésta, se puede apreciar en la figura 5.2 (b) que ya todas la estaciones cuentan con los valores en los tres años seleccionados, además se puede apreciar que la serie de tiempo es estacionaria, pues parece variar en torno a una media fija.



(a) Histograma de CO por estación



(b) Histograma de CO

FIGURA 5.3: Histogramas de CO

La figura 5.2 muestra los gráficos de series de tiempo para la variable CO por estación y la figura 5.3 muestra los histogramas de los datos por estación de la variable CO. De la figura 5.2 se puede observar que si bien las trece series de tiempo muestran características muy similares, es decir, son parecidas en estacionariedad en el tiempo, los histogramas de la figura 5.3 son diferentes entre algunas estaciones, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

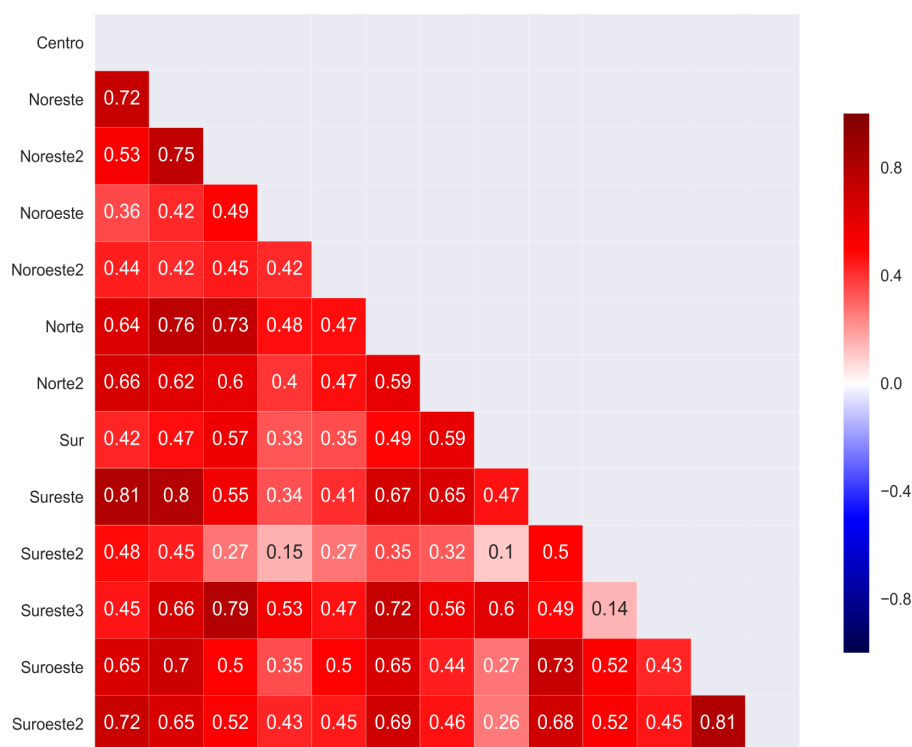


FIGURA 5.4: Matriz de correlación de CO entre estaciones

La figura 5.4 muestra que la variable CO se correlaciona positivamente alto (correlaciones mayores a 0.2) entre la mayoría de las estaciones, pues la estación Sureste 2 muestra correlaciones bajas (correlaciones menores a 0.2) contra las estaciones Noroeste, Sur y Sureste 3.

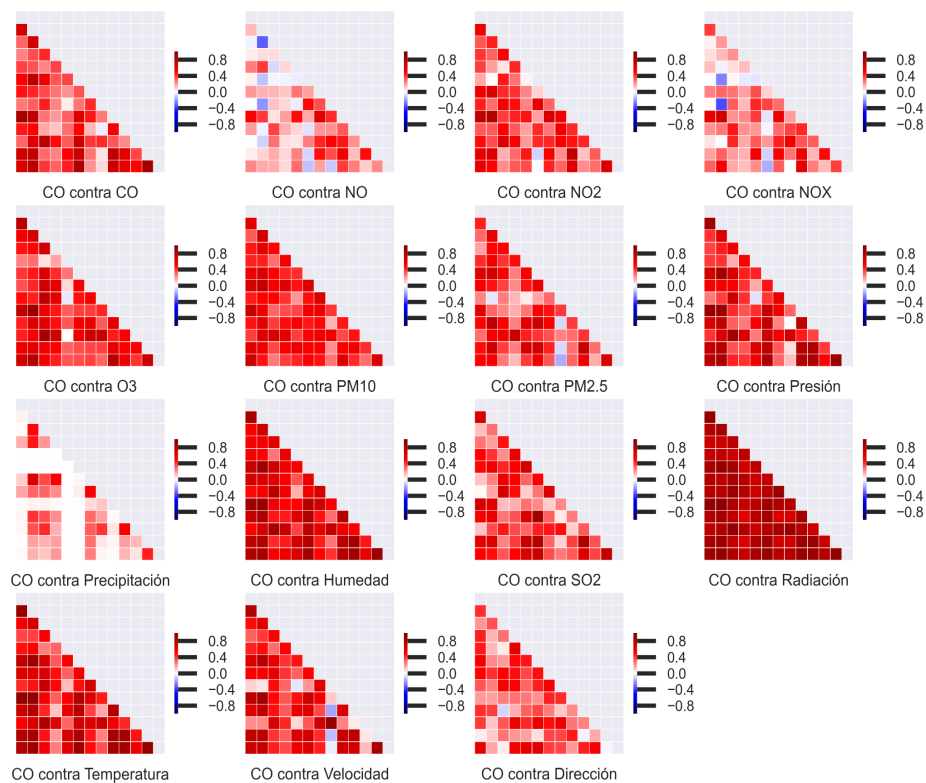
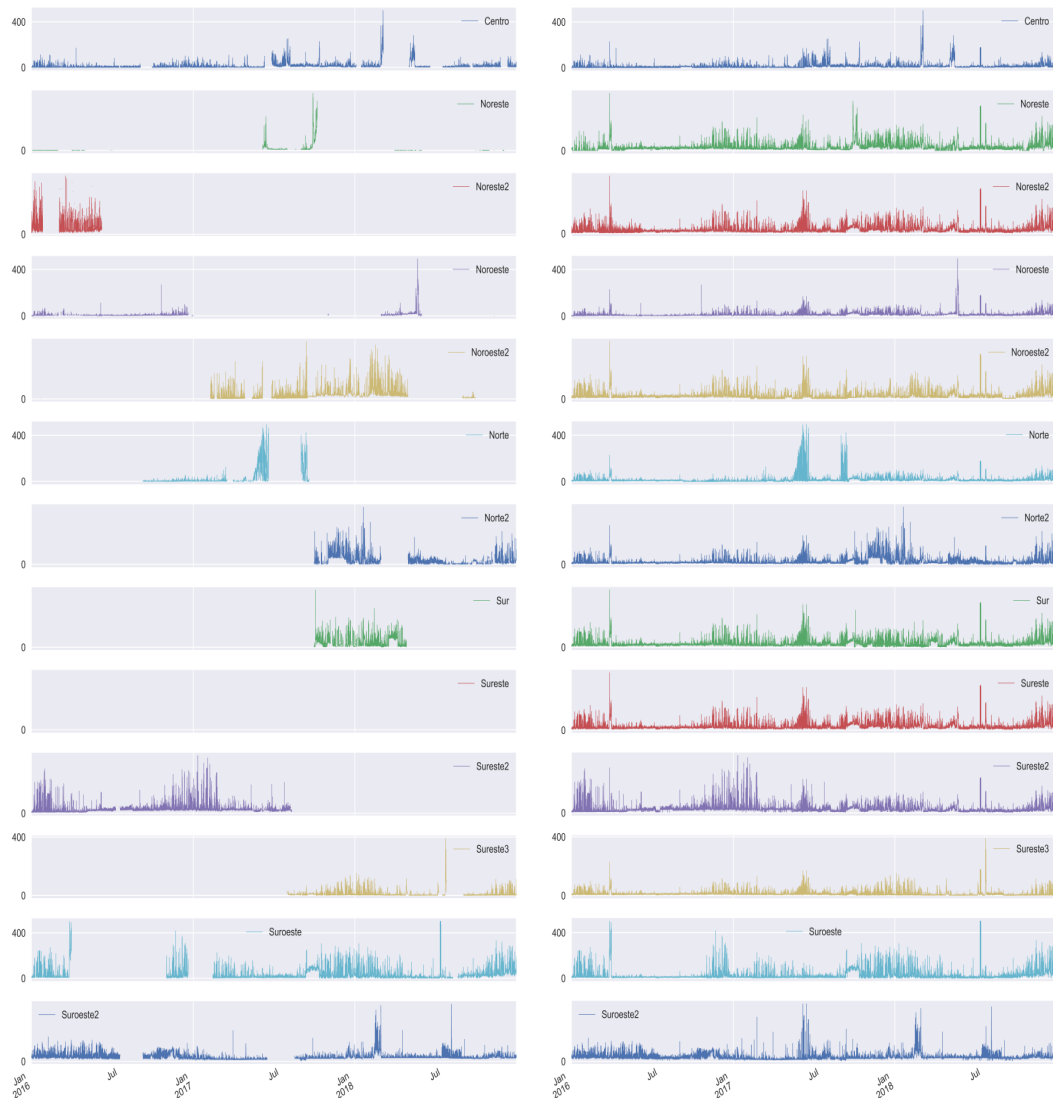


FIGURA 5.5: Matriz de correlación de CO contra el resto de las variables

La correlaciones de la variable CO con el resto de las variables se correlaciona positivamente mostrando correlaciones mayores a 0.2 (ver figura 5.5), sin embargo para la variable *precipitación pluvial* las correlaciones son menores a 0.2 (ver figura 5.5).

5.1.2 ÓXIDO NÍTRICO (NO)



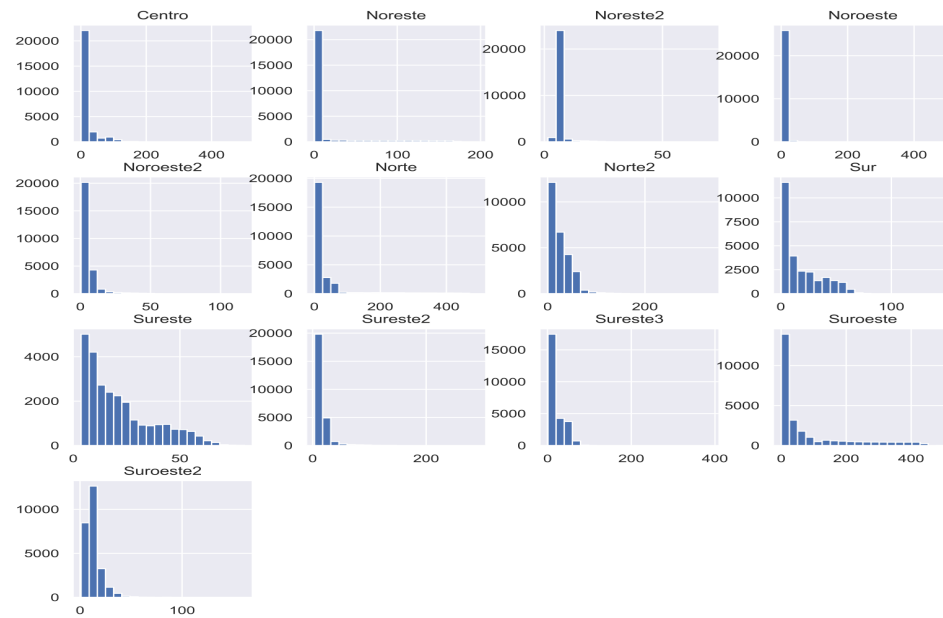
(a) Serie de tiempo de NO original

(b) Serie de tiempo de NO rellenando datos desconocidos

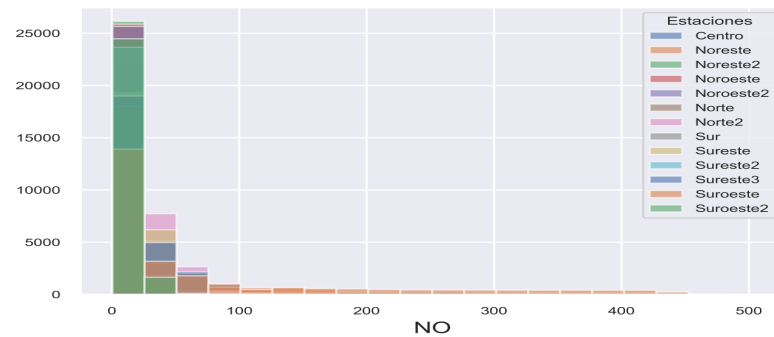
FIGURA 5.6: Series de tiempo de NO desde 2016 hasta 2018

La variable NO además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.6 que la serie de tiempo es estacionaria

pues parece variar entorno a una media fija.



(a) Histograma de NO por estación



(b) Histograma de NO

FIGURA 5.7: Histogramas de NO

La figura 5.6 muestra los gráficos de series de tiempo para la variable NO por estación y la figura 5.7 muestra un histograma de los datos por estación de la variable NO. De la figura 5.6 se puede observar que si bien las trece series de

tiempo muestran características muy similares, los histogramas de la figura 5.7 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

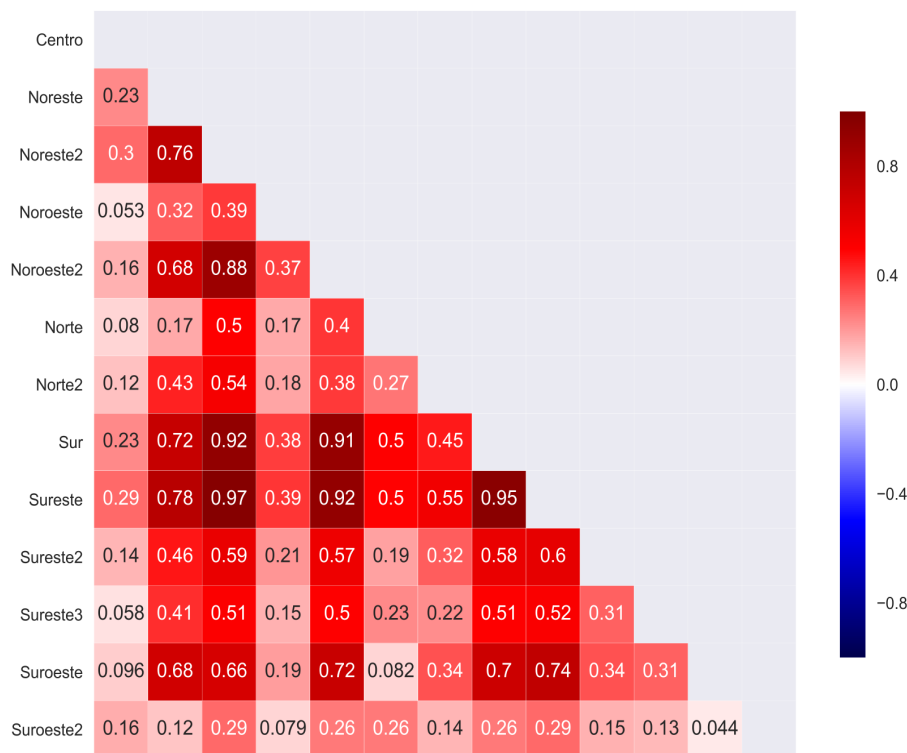


FIGURA 5.8: Matriz de correlación de NO entre estaciones

La figura 5.8 muestra que la variable NO se correlaciona positivamente alto (correlaciones mayores a 0.2) entre la mayoría de las estaciones, ya que las estaciones Centro y Sureste 2 muestran correlaciones menores a 0.2 contra el resto de la estaciones. Además para las estaciones Noroeste, Noreste y Norte muestran correlaciones bajas.

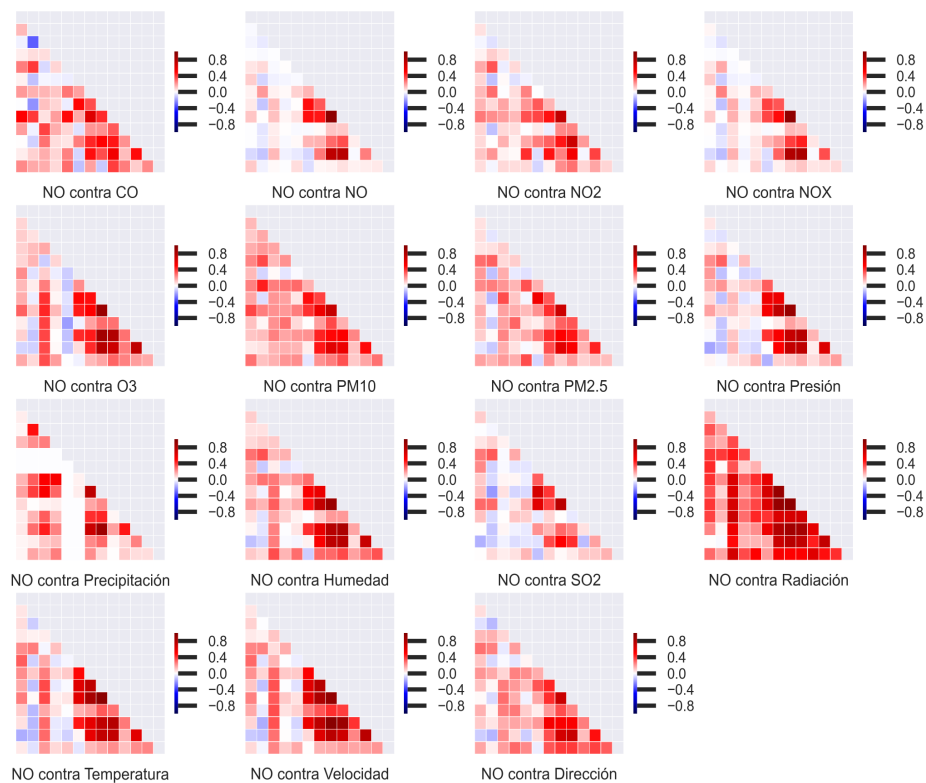


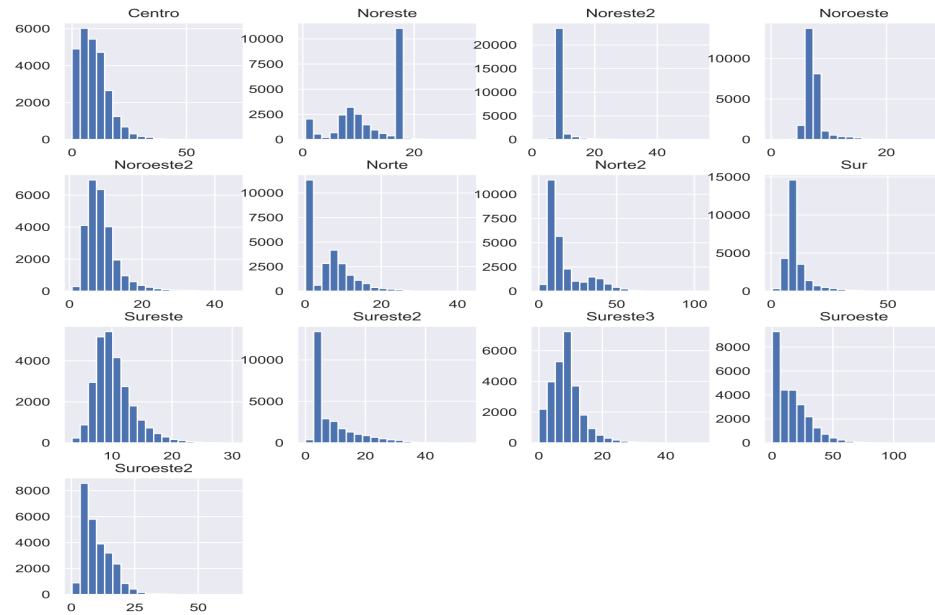
FIGURA 5.9: Matriz de correlación de NO contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* y NO_x , (ver figura 5.9) se presentan en algunas estaciones correlaciones altas y correlaciones nulas. Además, para las correlaciones con la variable NO_2 presenta correlaciones positivas pequeñas para la estación Suroeste 2.

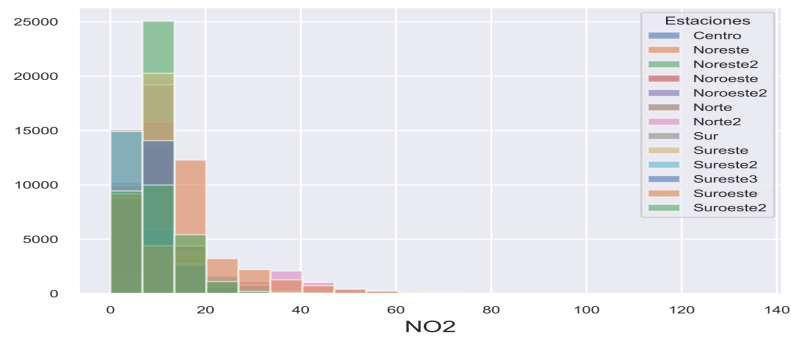
5.1.3 BIÓXIDO DE NITRÓGENO (NO_2)(a) Serie de tiempo de NO_2 original(b) Serie de tiempo de NO_2 rellenando datos desconocidosFIGURA 5.10: Series de tiempo de NO_2 desde 2016 hasta 2018

La variable NO_2 presenta datos faltantes durante periodos de tiempo en los tres años que se rellenan haciendo el promedio de las estaciones cercanas, este promedio se le asigna a las estaciones sin registros. También, se puede apreciar

en la figura 5.10 (b), que la serie de tiempo es estacionaria, ya que parece variar entorno a una media, presentando altas variaciones en el mes de julio del 2017 en las trece estaciones.



(a) Histograma de NO_2 por estación



(b) Histograma de NO_2

FIGURA 5.11: Histogramas de NO_2

La figura 5.10 muestra los gráficos de series de tiempo para la variable NO_2 por estación y la figura 5.11 muestra los histogramas de los datos por estación de

la variable NO_2 . En la figura 5.10 se puede observar que si bien las trece series de tiempo muestran características muy similares, también, los histogramas de la figura 5.11 son muy similares a excepción de las estaciones Norte 2 y Sur.

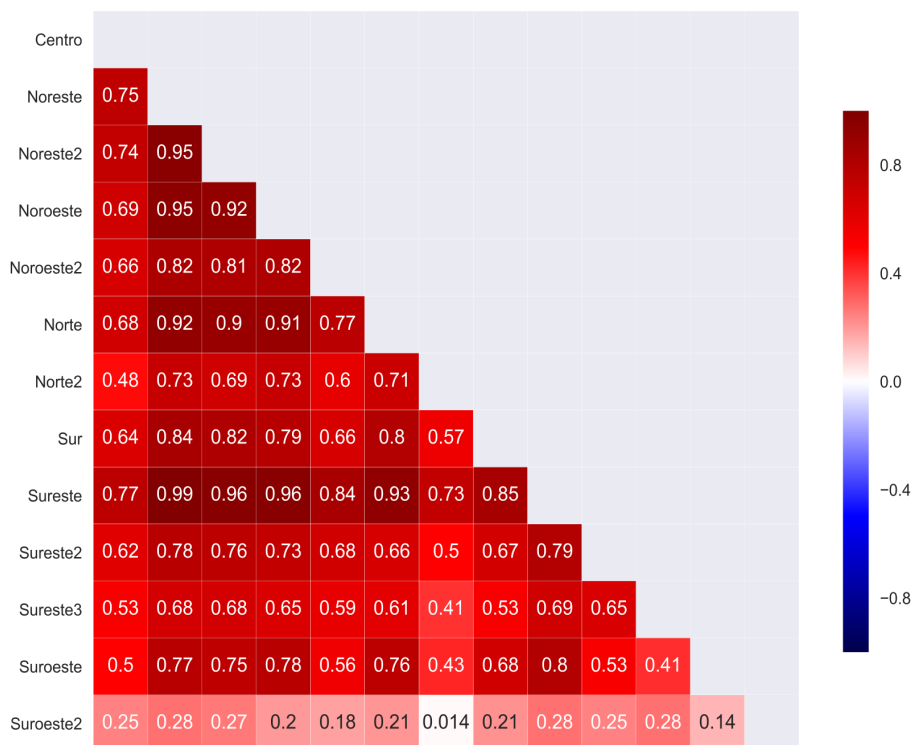


FIGURA 5.12: Matriz de correlación de NO_2 entre estaciones

La figura 5.12 muestra que la variable NO_2 se correlaciona positivamente entre la mayoría de las estaciones, ya que la estación Suroeste 2 muestra correlaciones bajas contra las estaciones Noroeste 2, Norte 2 y Suroeste.

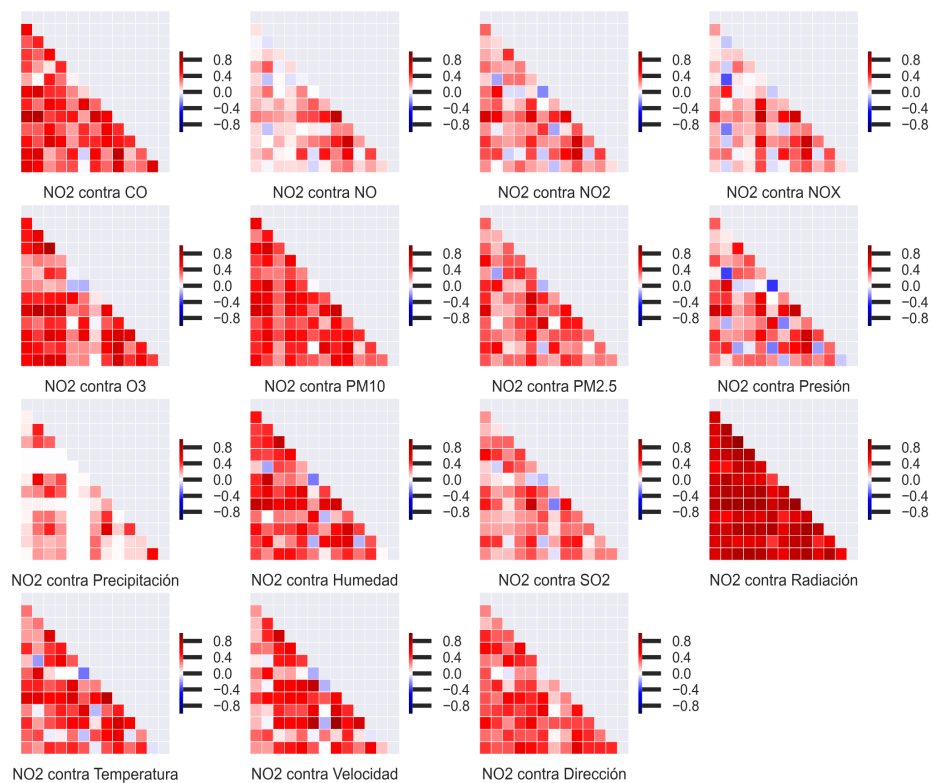
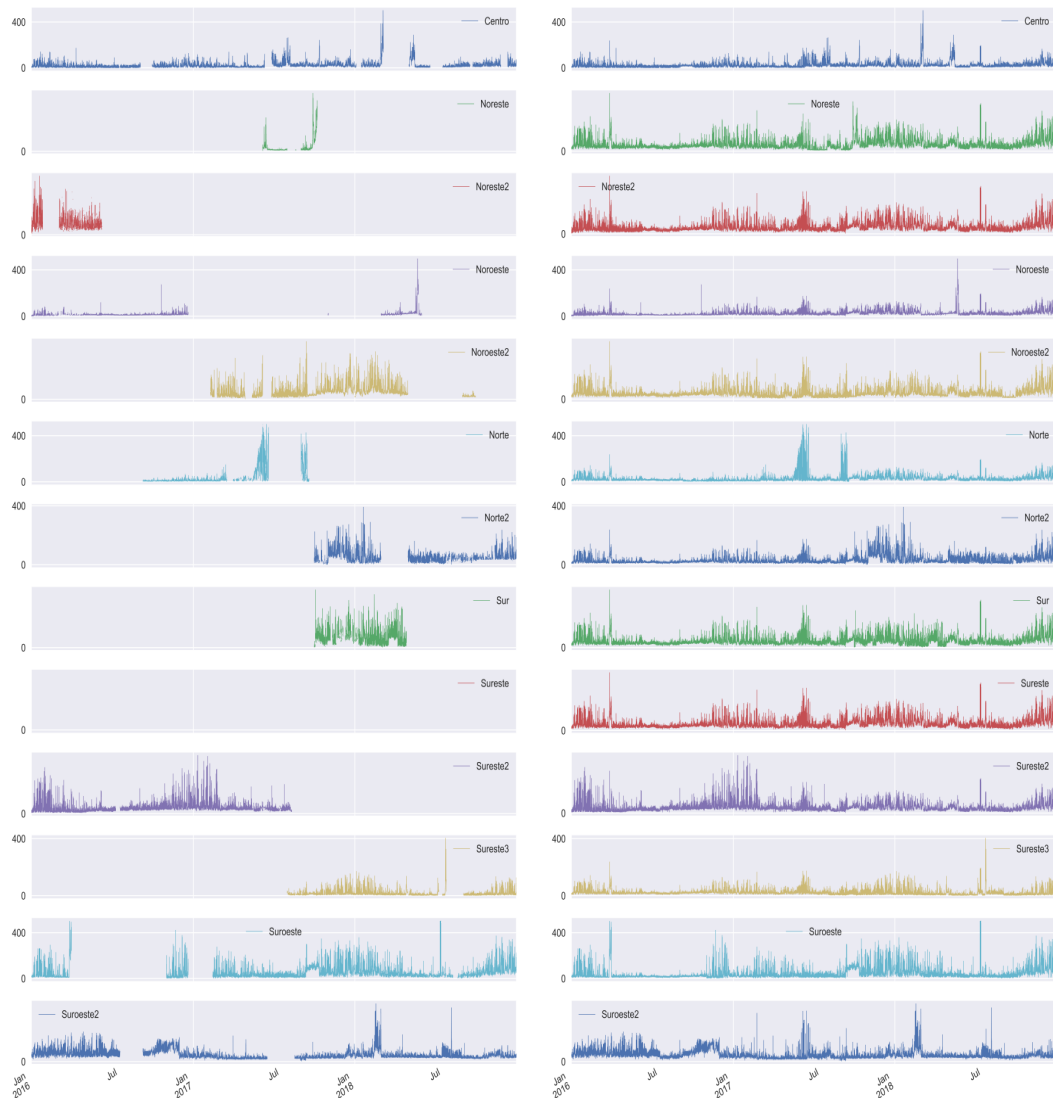


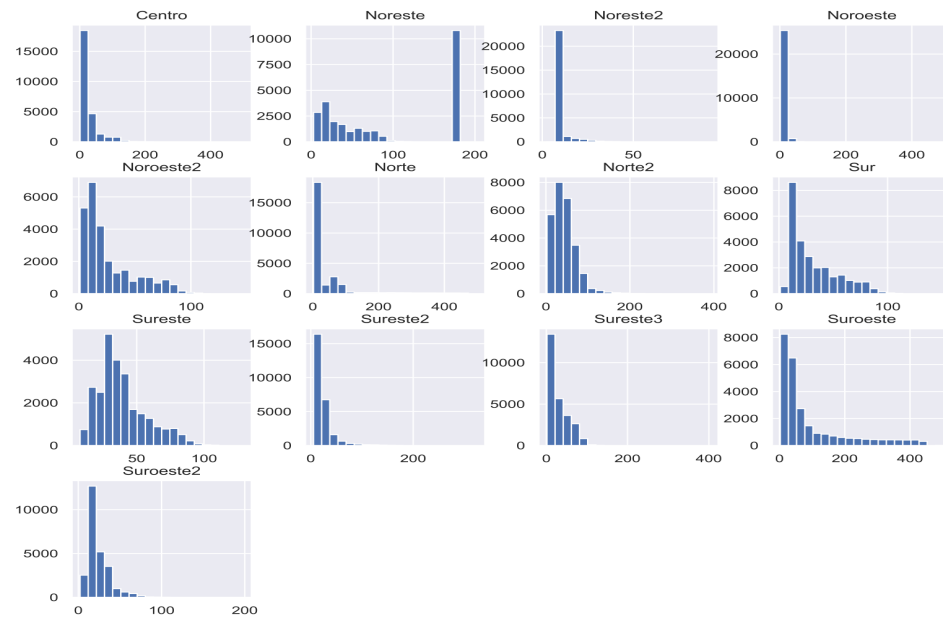
FIGURA 5.13: Matriz de correlación de NO₂ contra el resto de las variables

La estación Suroeste 2 presenta correlaciones positivas pequeñas contra la variable NO_x (ver figura 5.13). Para las correlaciones de la variable *precipitación pluvial* se presentan en algunas estaciones correlaciones altas y correlaciones cercanas a cero (ver figura 5.13).

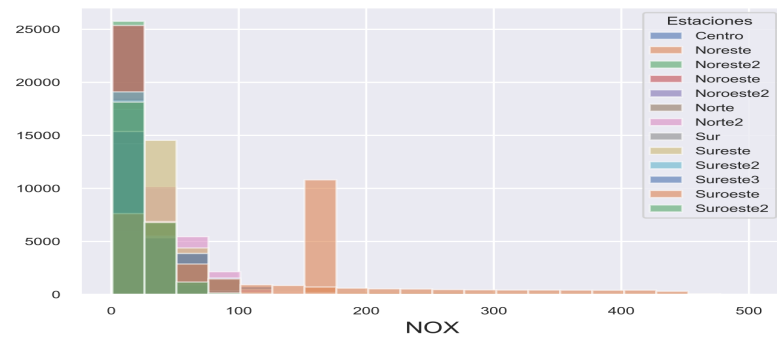
5.1.4 ÓXIDOS DE NITRÓGENO (NO_X)(a) Serie de tiempo de NO_X original(b) Serie de tiempo de NO_X rellenando datos desconocidosFIGURA 5.14: Series de tiempo de NO_X desde 2016 hasta 2018

La variable NO_X además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.14 que la serie de tiempo es

estacionaria ya que parece variar en torno a una media fija.



(a) Histograma de NO_x por estación



(b) Histograma de NO_x

FIGURA 5.15: Histogramas de NO_x

La figura 5.14 muestra los gráficos de series de tiempo para la variable NO_x por estación y la figura 5.15 muestra un histograma de los datos por estación de la variable NO_x . De la figura 5.14 se puede apreciar que en las trece series de

tiempo muestran características muy similares, sin embargo, los histogramas de la figura 5.15 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacer esto, se pierden las características importantes de los datos que dependen del tiempo.

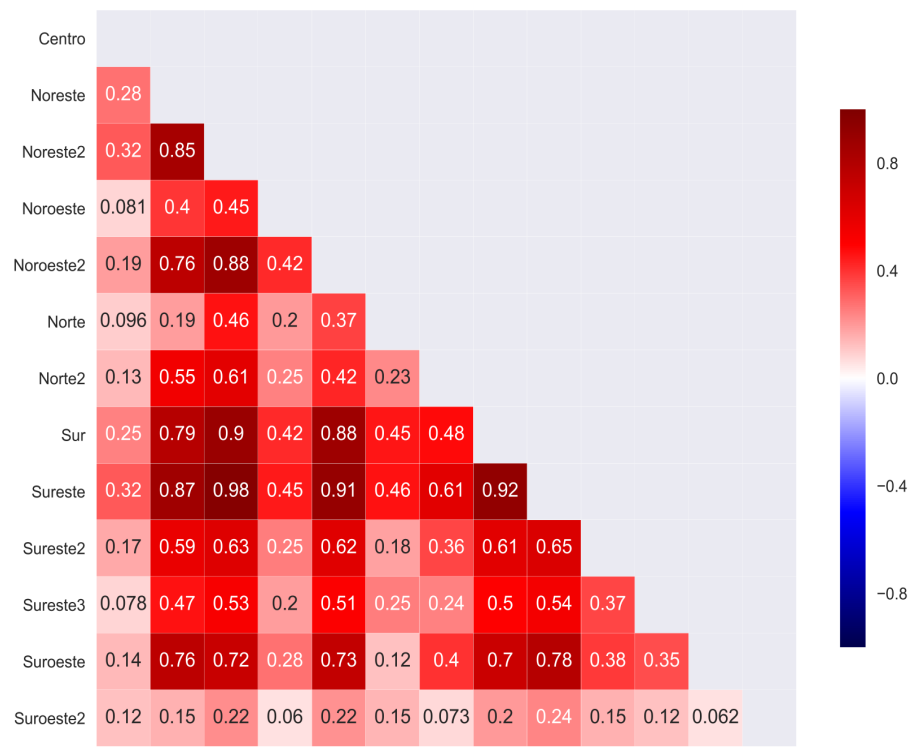


FIGURA 5.16: Matriz de correlación de NO_x entre estaciones

La correlación de la variable NO_x entre estaciones (NO_x vs NO_x) presenta correlaciones positivas pequeñas (ver figura 5.16).

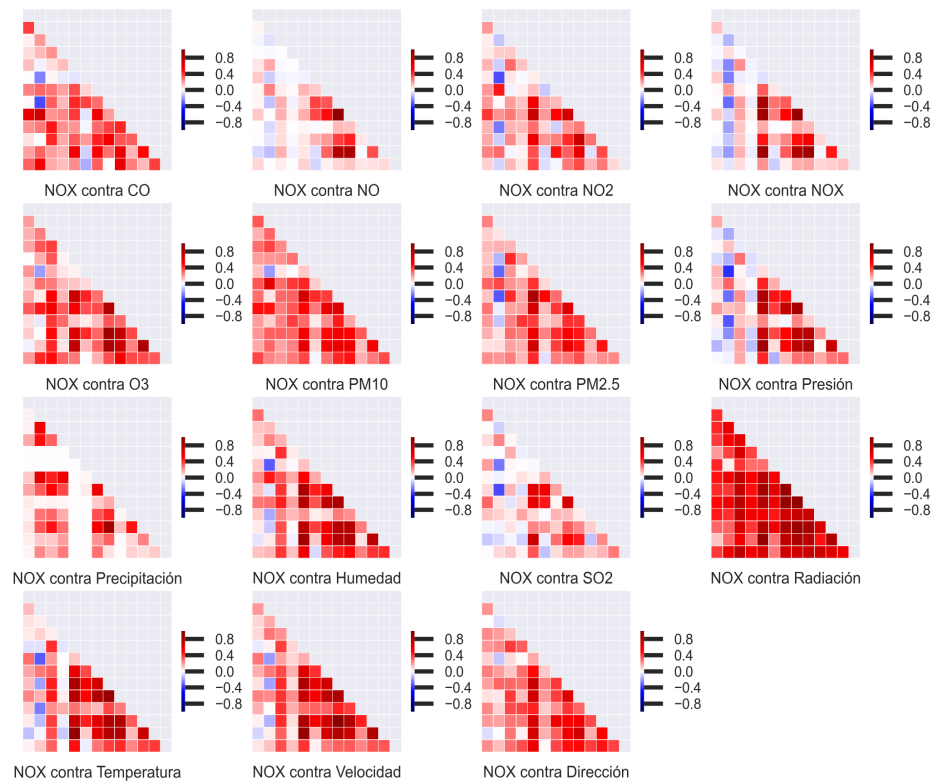
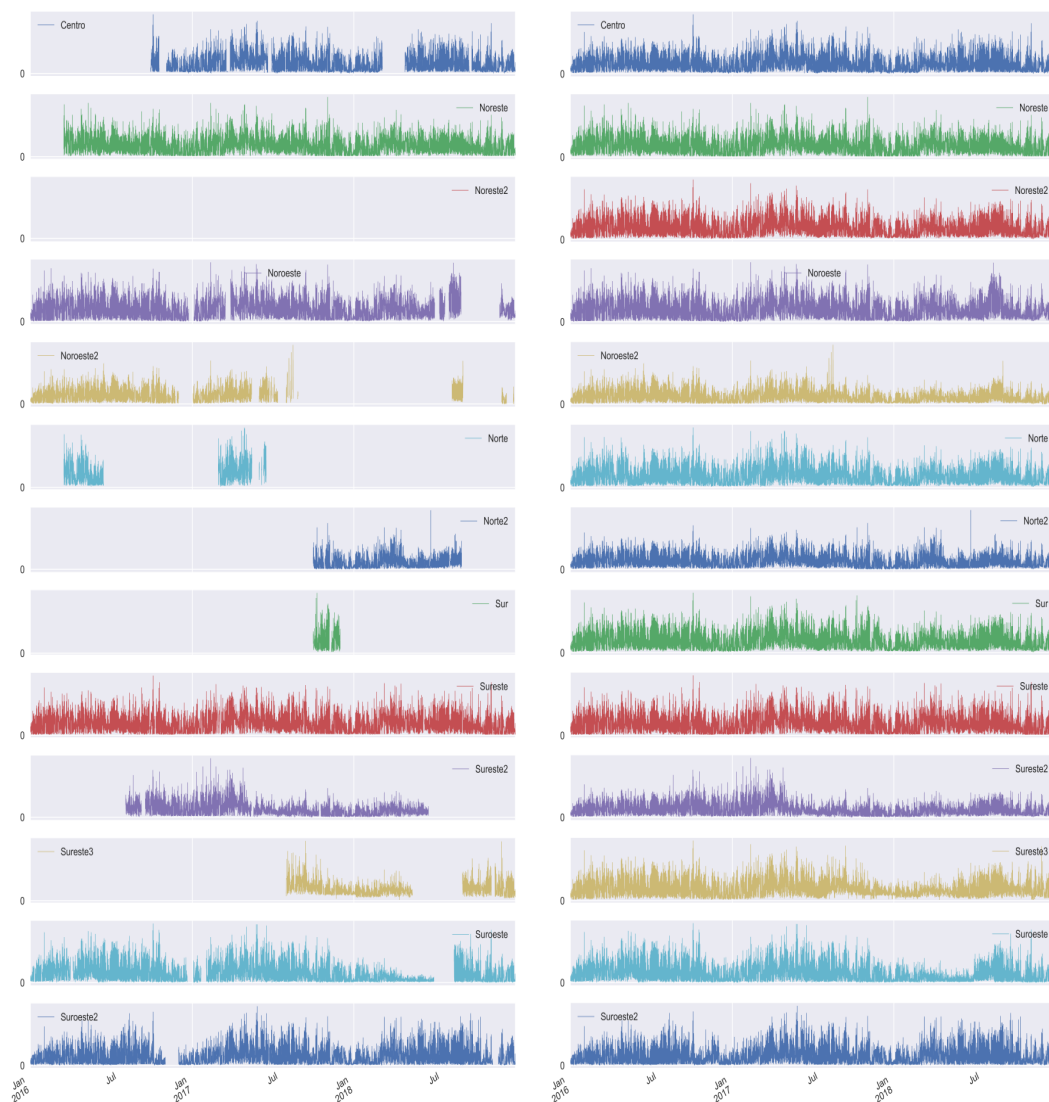


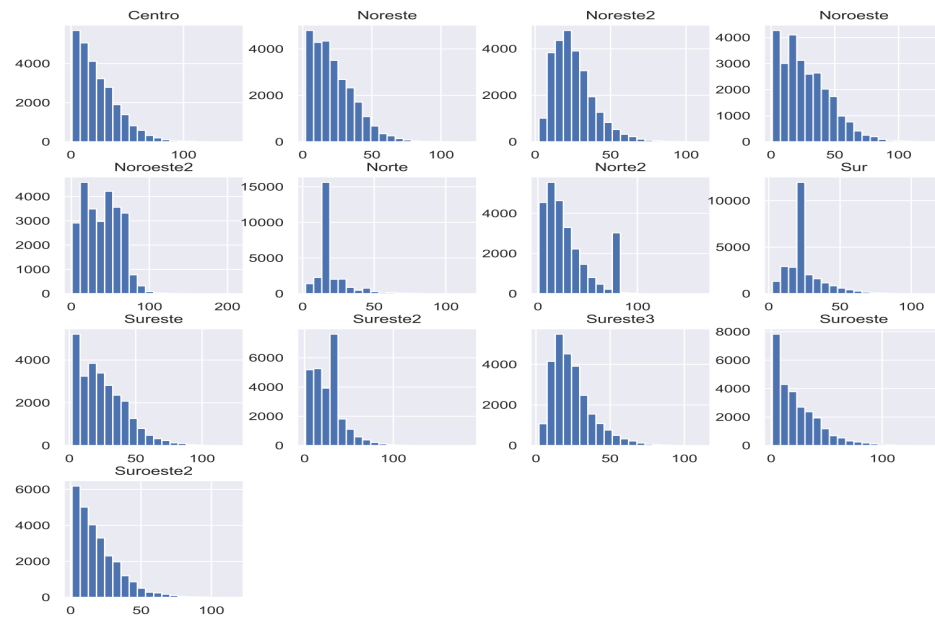
FIGURA 5.17: Matriz de correlación de NO_x contra el resto de las variables

Las correlaciones de la variable *precipitación pluvial* (ver Figura 5.17) muestran que para algunas correlaciones entre estaciones se presentan correlaciones positivas bajas.

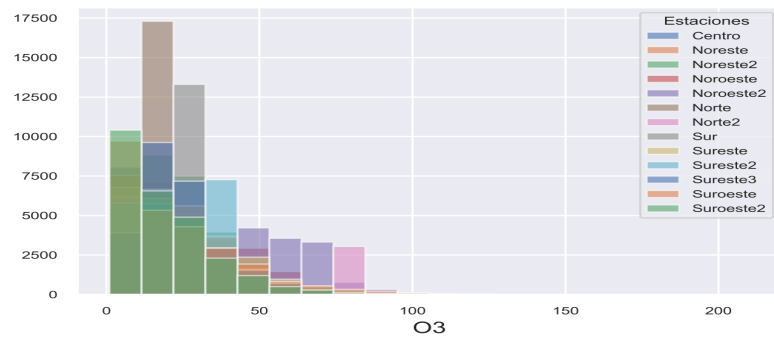
5.1.5 OZONO (O_3)(a) Serie de tiempo de O_3 original(b) Serie de tiempo de O_3 rellenando datos desconocidosFIGURA 5.18: Series de tiempo de O_3 desde 2016 hasta 2018

La variable O_3 además de tener datos faltantes que se completan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.18 que la serie de tiempo es

estacionaria pues varía entorno a un valor.



(a) Histograma de O_3 por estación



(b) Histograma de O_3

FIGURA 5.19: Histogramas de O_3

La figura 5.18 muestra los gráficos de series de tiempo para la variable O_3 por estación y la figura 5.19 muestra un histograma de los datos por estación de la variable O_3 . De la figura 5.18 se puede observar que si bien las trece series

de tiempo muestran características similares, los histogramas de la figura 5.19 son distintos, pues el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características principales de los datos que dependen del tiempo.

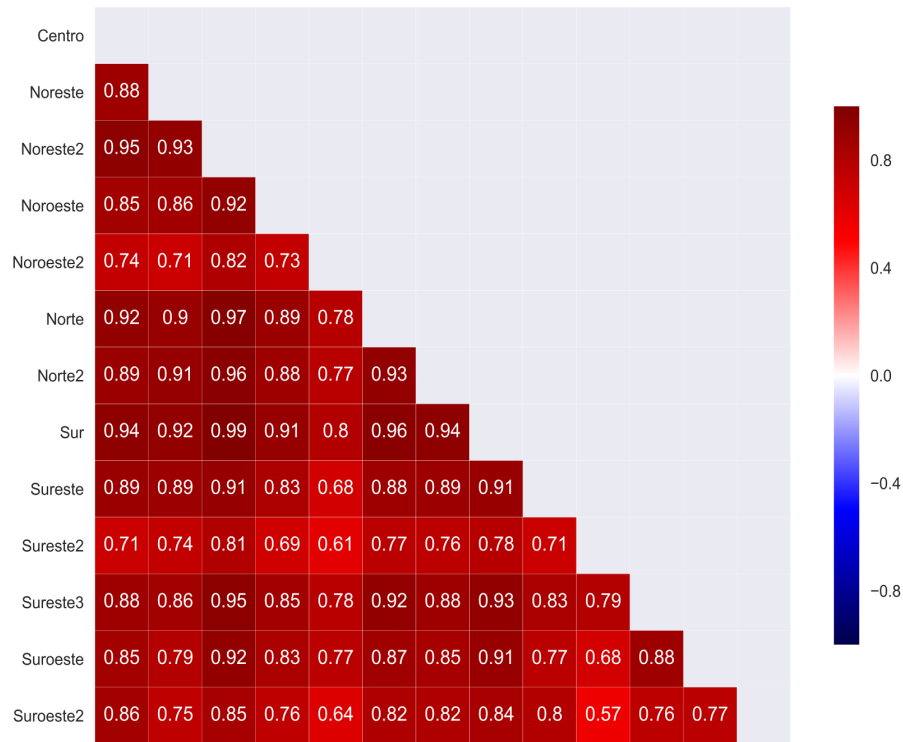


FIGURA 5.20: Matriz de correlación de O_3 entre estaciones

La figura 5.20 muestra que la variable O_3 se correlaciona positivamente alto en todas las estaciones ya que la mayor correlación es de 0.99 entre las estaciones Noroeste 2 y Sur, mientras que la menor correlación encontrada es de 0.57 entre las estaciones Suroeste 2 y Sureste 2.

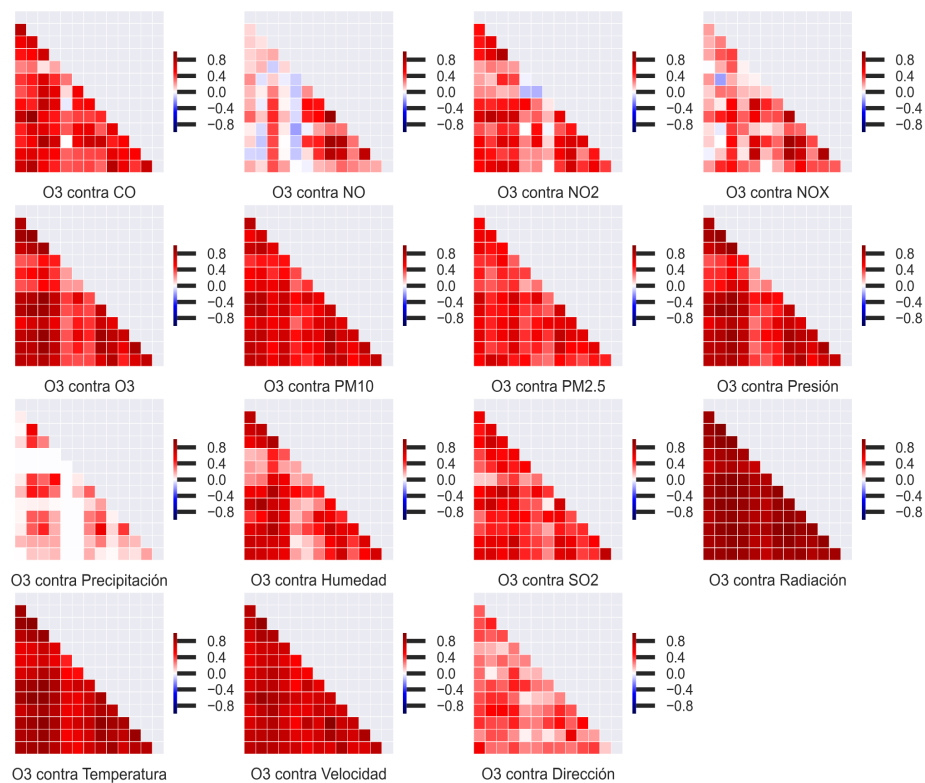
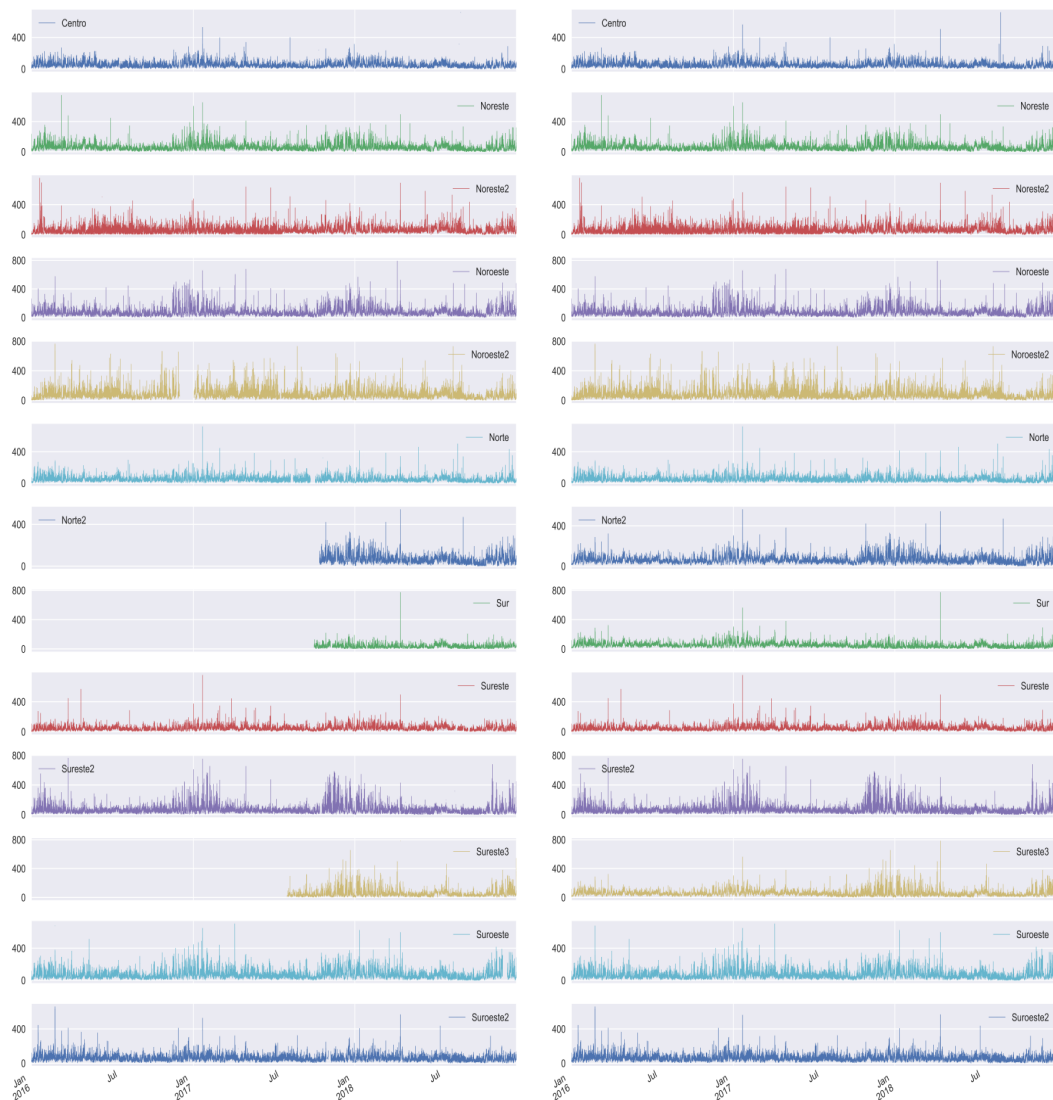


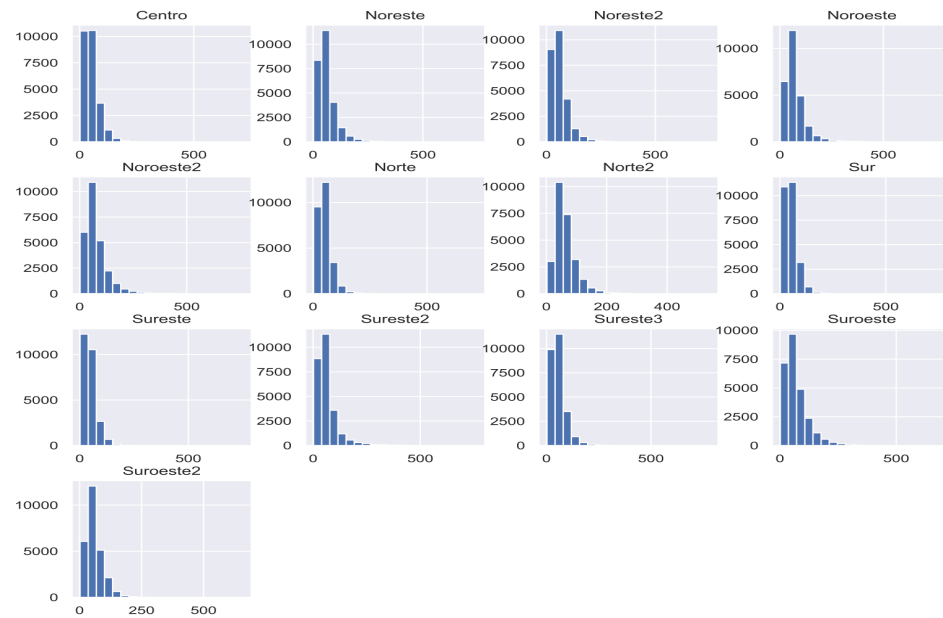
FIGURA 5.21: Matriz de correlación de O₃ contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* se presentan en algunas estaciones correlaciones altas y correlaciones cercanas a cero (ver figura 5.21), para el resto de las variables la variable O₃ se correlaciona positivamente alto en todas las estaciones.

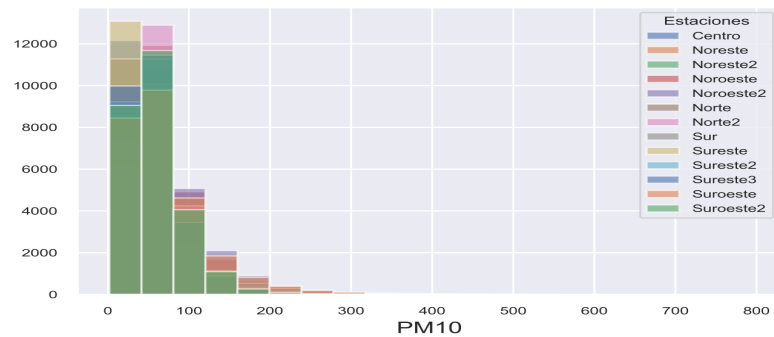
5.1.6 PARTÍCULAS MENORES A 10 MICRAS (PM_{10})(a) Serie de tiempo de PM_{10} original(b) Serie de tiempo de PM_{10} rellenando datos desconocidosFIGURA 5.22: Series de tiempo de PM_{10} desde 2016 hasta 2018

La variable PM_{10} además de tener datos faltantes que se completan haciendo el promedio de las estaciones cercanas y luego se asigna el promedio a las estaciones sin datos, se puede apreciar en la figura 5.22 que la serie de tiempo

es estacionaria pues varía entorno a una media.



(a) Histograma de PM_{10} por estación



(b) Histograma de PM_{10}

FIGURA 5.23: Histogramas de PM_{10}

La figura 5.22 muestra los gráficos de series de tiempo para la variable PM_{10} por estación y la figura 5.23 muestra un histograma de los datos por estación de la variable PM_{10} . De la figura 5.22 se puede observar que las trece series de tiempo

muestran características similares, pero, los histogramas de la figura 5.23 son diferentes, pues el histograma resume los datos a través de la dimensión del tiempo, y al hacer esto, se pierden las características principales de los datos que dependen del tiempo.

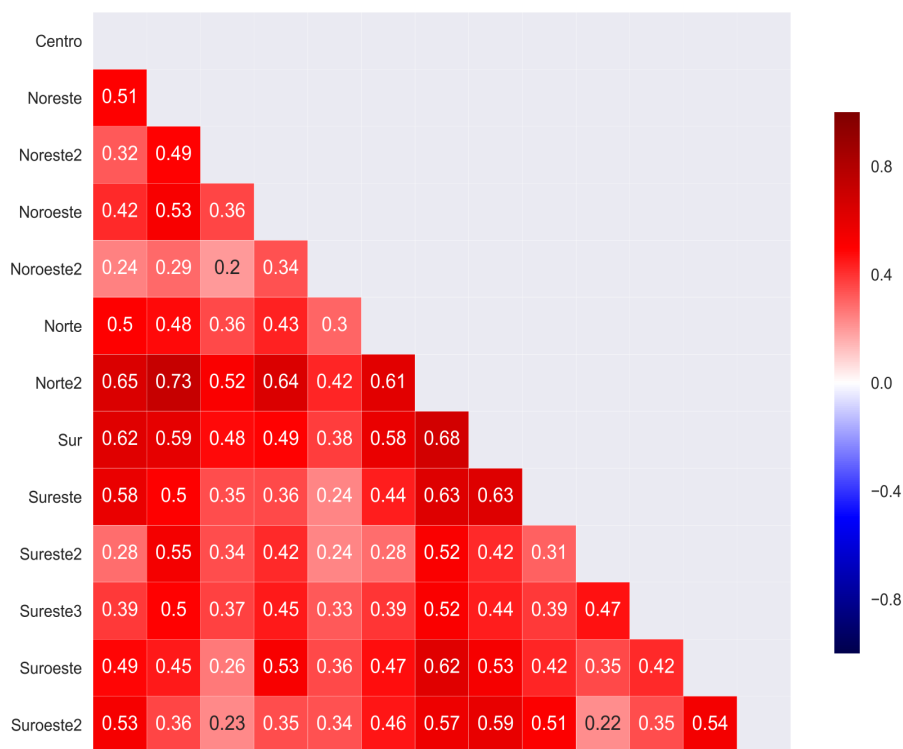


FIGURA 5.24: Matriz de correlación de PM_{10} entre estaciones

La figura 5.24 muestra que la variable PM_{10} se correlaciona positivamente alto entre todas las estaciones ya que la mayor correlación es de 0.73 entre las estaciones Norte 2 y Noreste, mientras que la menor correlación encontrada es de 0.20 entre las estaciones Noreste 2 y Noroeste 2.

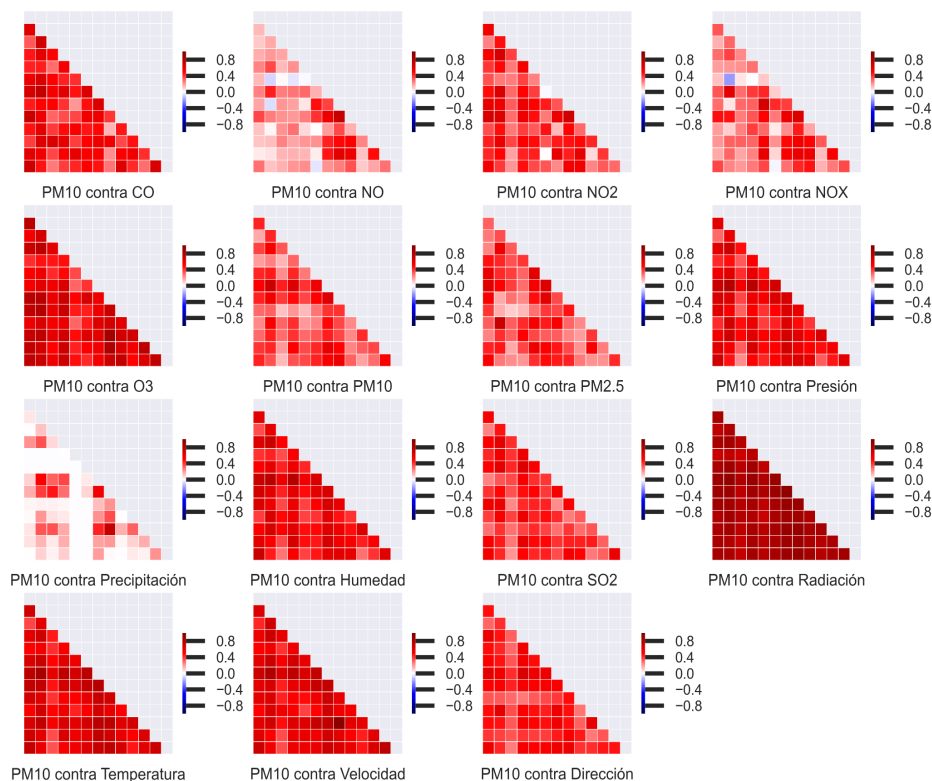
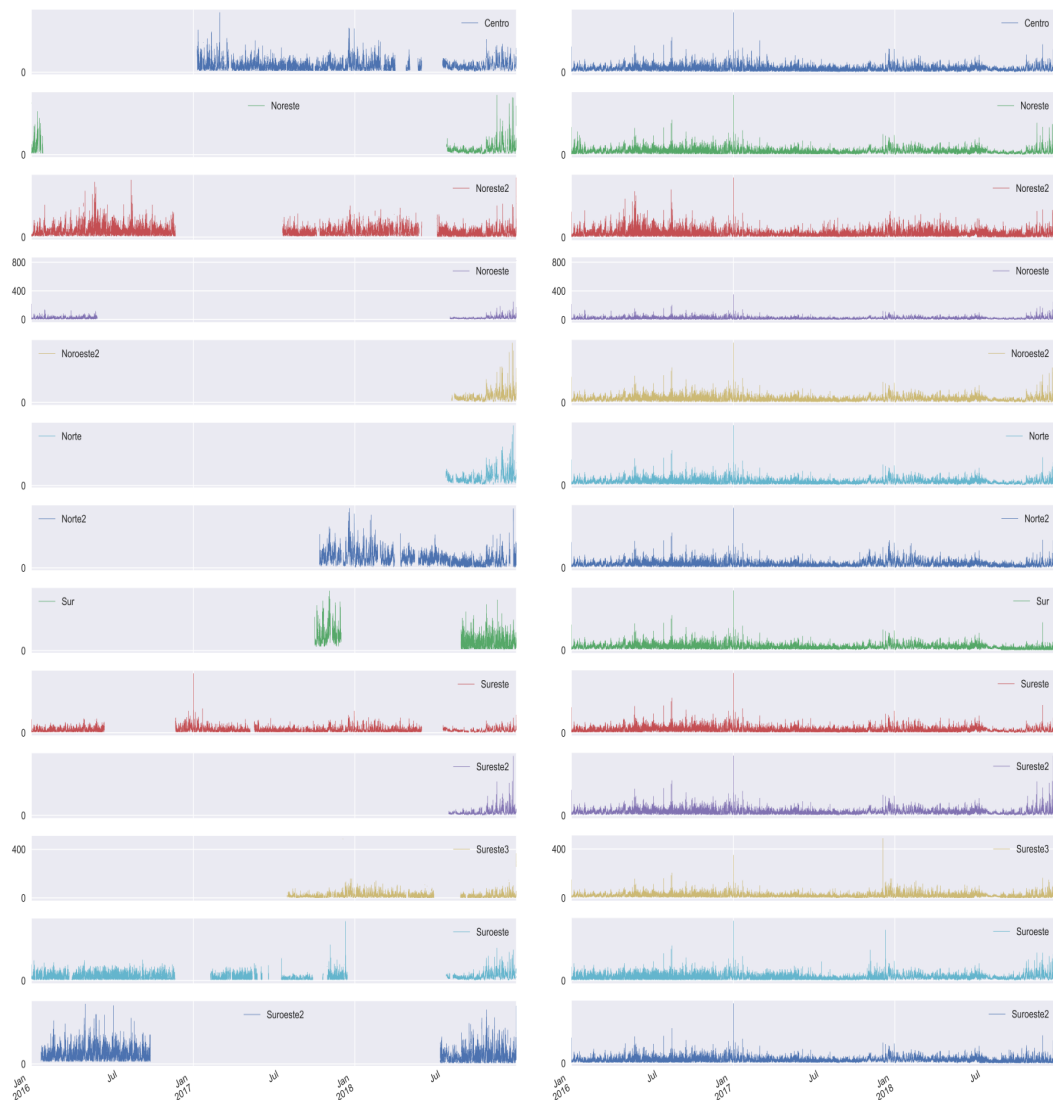


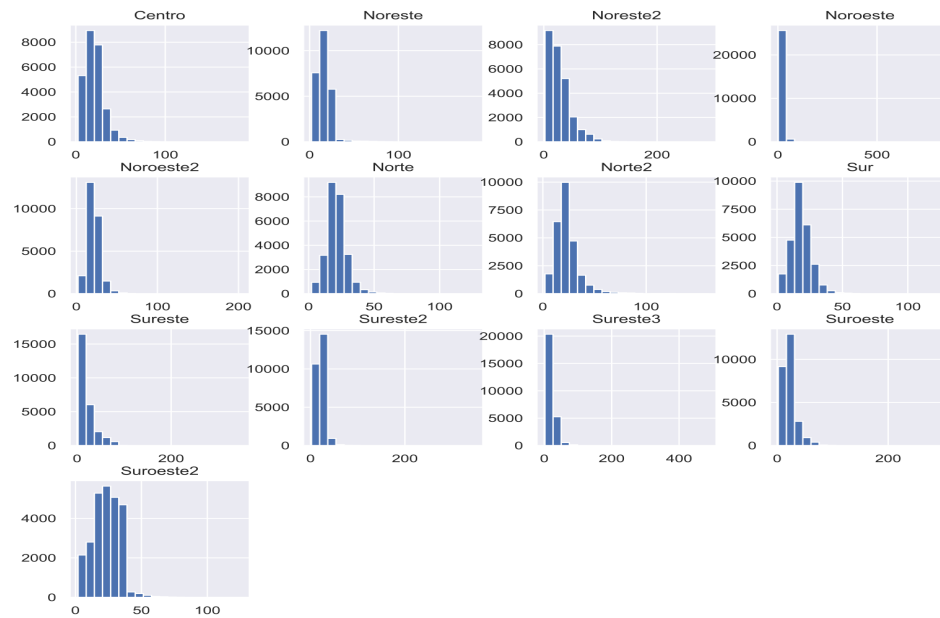
FIGURA 5.25: Matriz de correlación de PM_{10} contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* se presentan algunas estaciones con correlaciones altas y correlaciones cercanas a cero (ver figura 5.25), para el resto de las variables la variable PM_{10} se correlaciona positivamente alto en todas las estaciones.

5.1.7 PARTÍCULAS MENORES A 2.5 MICRAS ($PM_{2.5}$)(a) Serie de tiempo de $PM_{2.5}$ original(b) Serie de tiempo de $PM_{2.5}$ rellenando datos desconocidosFIGURA 5.26: Series de tiempo de $PM_{2.5}$ desde 2016 hasta 2018

La variable $PM_{2.5}$ además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.26 que la serie de tiempo es

estacionaria ya que parece variar en torno a una media fija.



(a) Histograma de $PM_{2.5}$ por estación



(b) Histograma de $PM_{2.5}$

FIGURA 5.27: Histogramas de $PM_{2.5}$

La figura 5.26 muestra los gráficos de series de tiempo para la variable $PM_{2.5}$ por estación y la figura 5.27 muestra un histograma de los datos por estación de la variable $PM_{2.5}$. De la figura 5.26 se puede observar que si bien las trece

series de tiempo muestran características muy similares, los histogramas de la figura 5.27 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

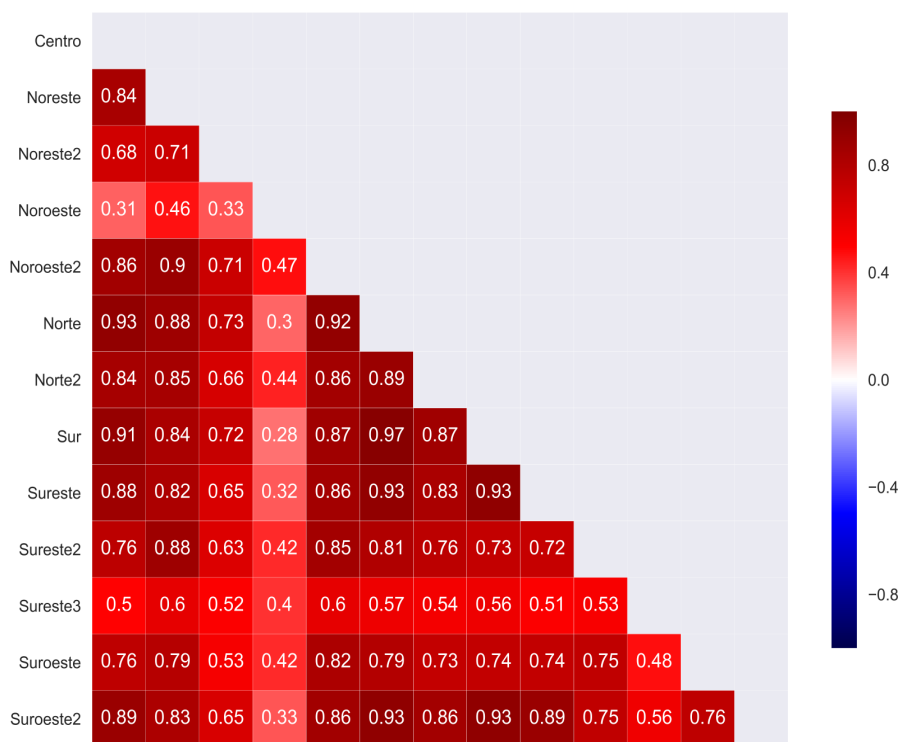


FIGURA 5.28: Matriz de correlación de $PM_{2.5}$ entre estaciones

La figura 5.28 muestra que la variable $PM_{2.5}$ se correlaciona positivamente alto entre todas las estaciones ya que la mayor correlación es de 0.97 entre las estaciones Norte y Sur, mientras que la menor correlación encontrada es de 0.28 entre las estaciones Sur y Noroeste.

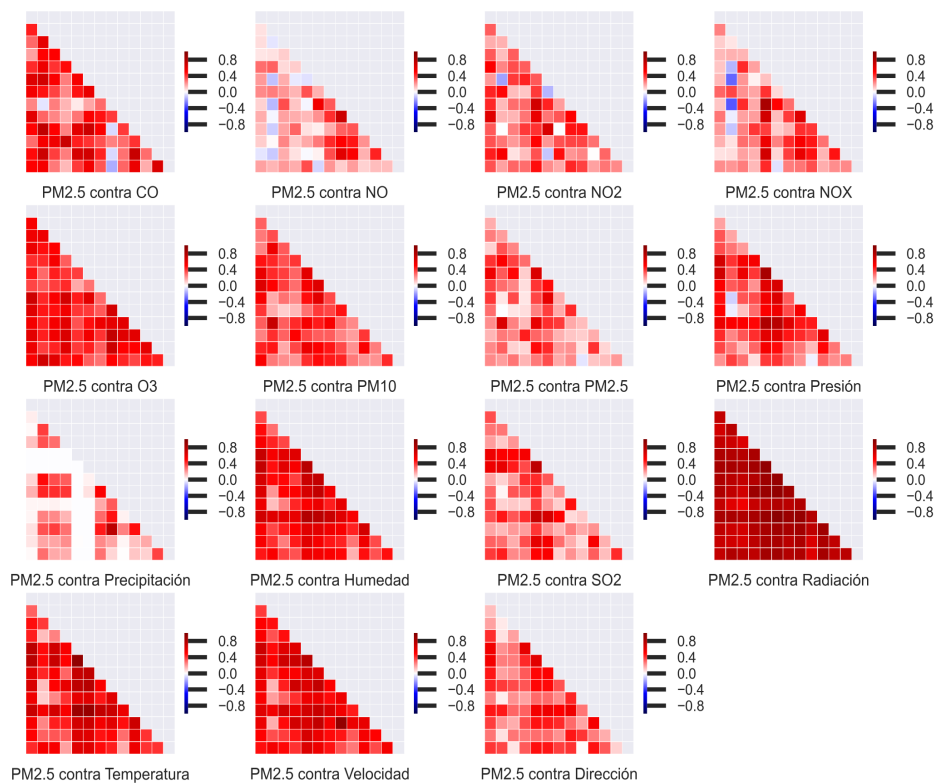
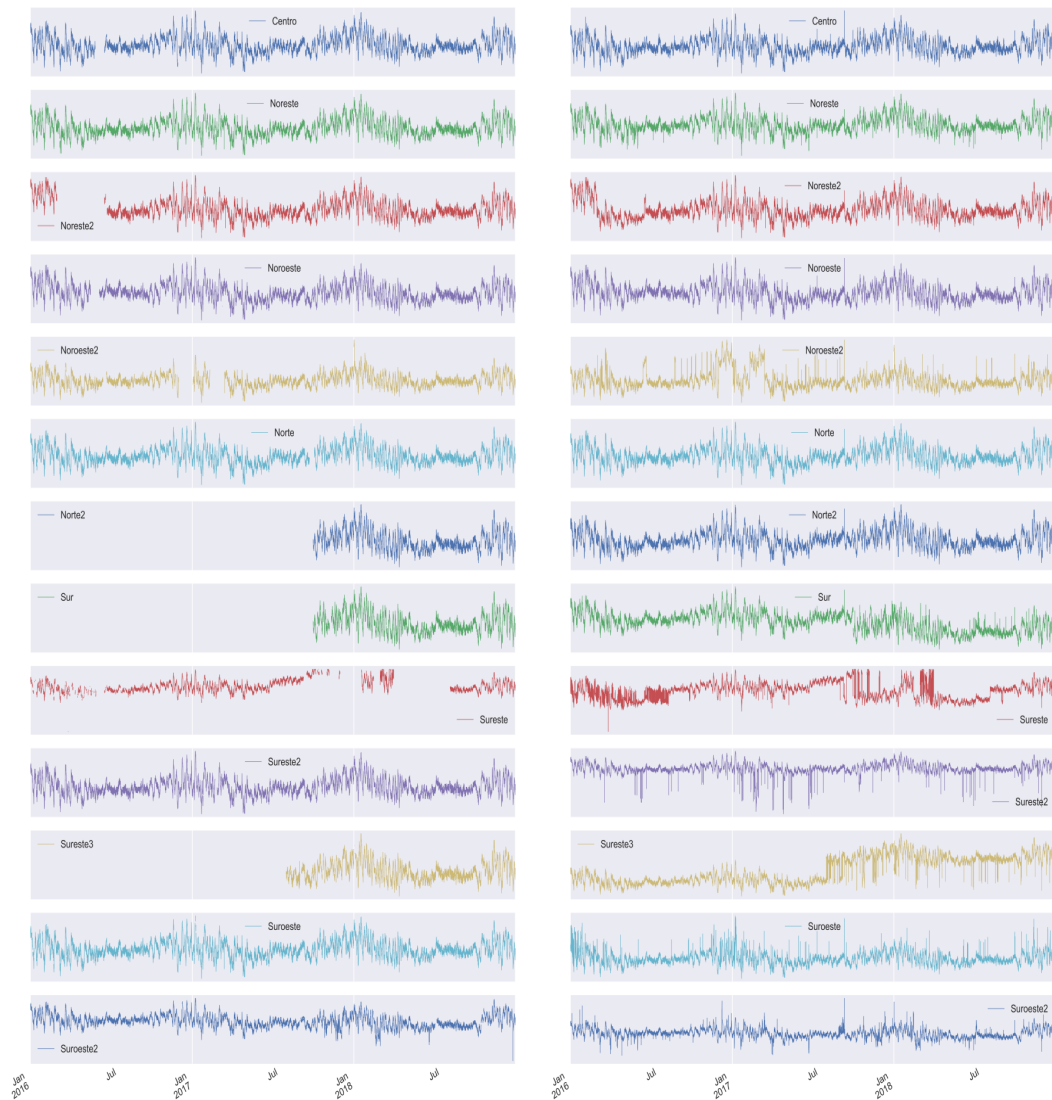


FIGURA 5.29: Matriz de correlación de PM_{2.5} contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* se presentan en algunas estaciones correlaciones altas y correlaciones cercanas a cero (ver figura 5.29), para el resto de las variables la variable PM_{2.5} se correlaciona positivamente alto en todas las estaciones.

5.1.8 PRESIÓN BAROMÉTRICA

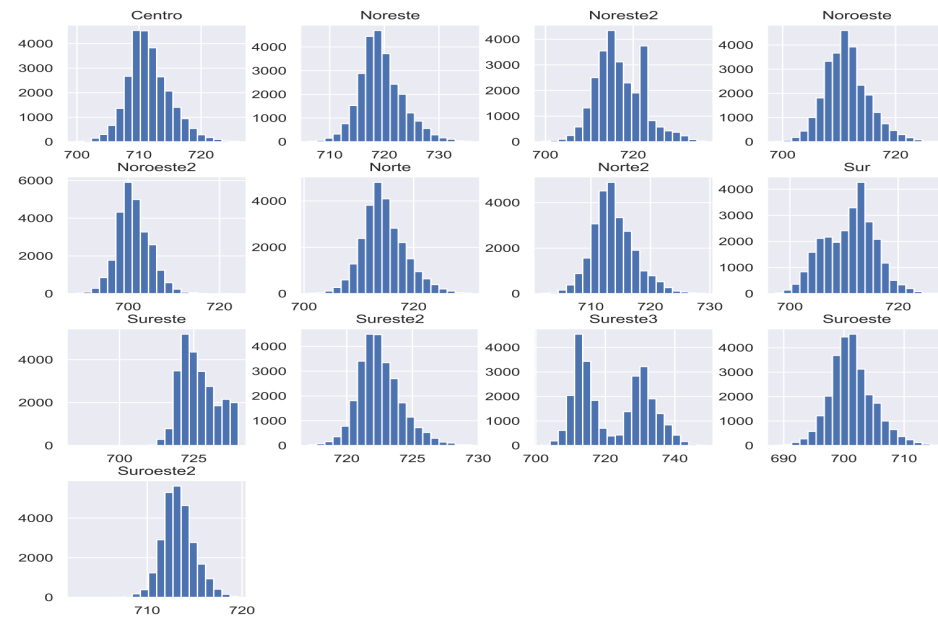


(a) Serie de tiempo de *presión barométrica* original
 (b) Serie de tiempo de *presión barométrica* relleno de datos desconocidos

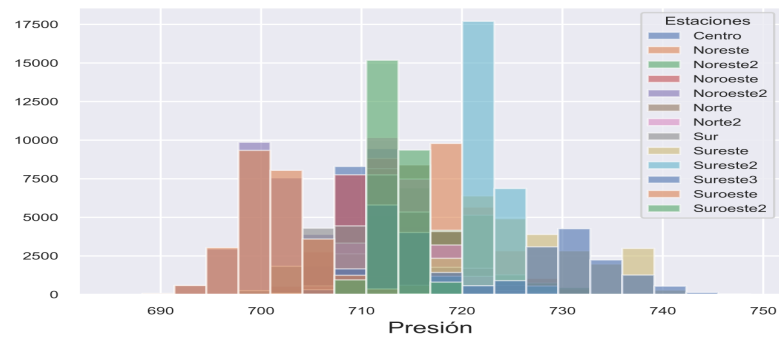
FIGURA 5.30: Series de tiempo de *presión barométrica* desde 2016 hasta 2018

La variable *presión barométrica* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.30 que la serie de tiempo

es estacionaria ya que parece variar entorno a una media.



(a) Histograma de *presión barométrica* por estación



(b) Histograma de *presión barométrica*

FIGURA 5.31: Histogramas de *presión barométrica*

La figura 5.30 muestra los gráficos de series de tiempo para la variable *presión barométrica* por estación y la figura 5.35 muestra un histograma de los datos por estación de la variable *presión barométrica*. De la figura 5.30 se puede

observar que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.35 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características más importantes de los datos que dependen del tiempo.

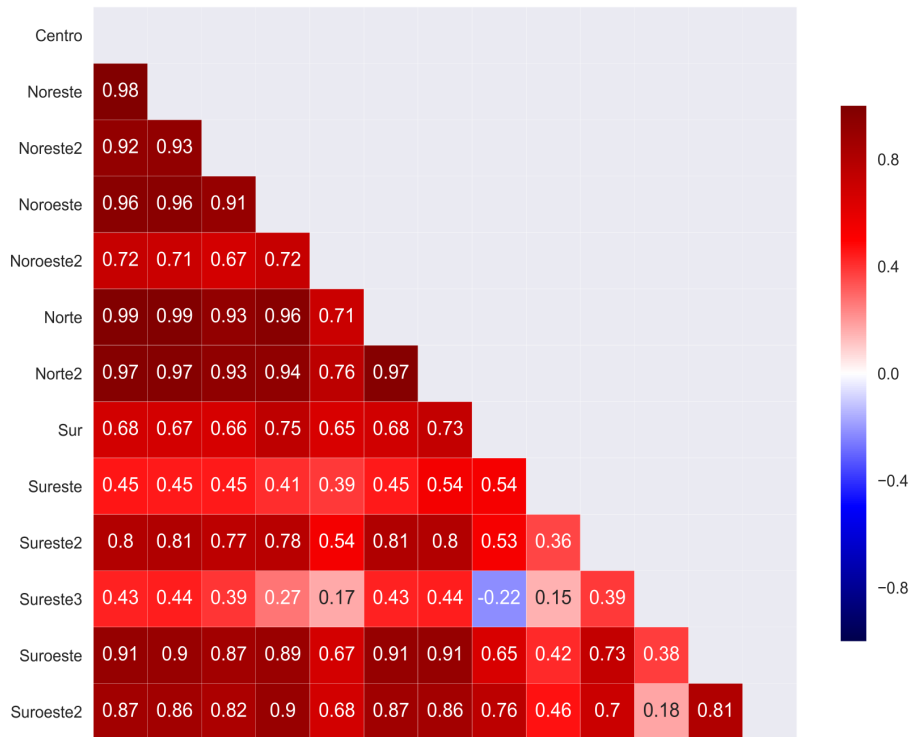


FIGURA 5.32: Matriz de correlación de *presión barométrica* entre estaciones

La figura 5.32 muestra que la variable *presión barométrica* se correlaciona positivamente alto entre la mayoría de las estaciones ya que la estación Sureste 3 presenta dos correlaciones positivas bajas contra las estaciones Noroeste 2 y Sureste y una correlación negativa alta contra la estación Sur.

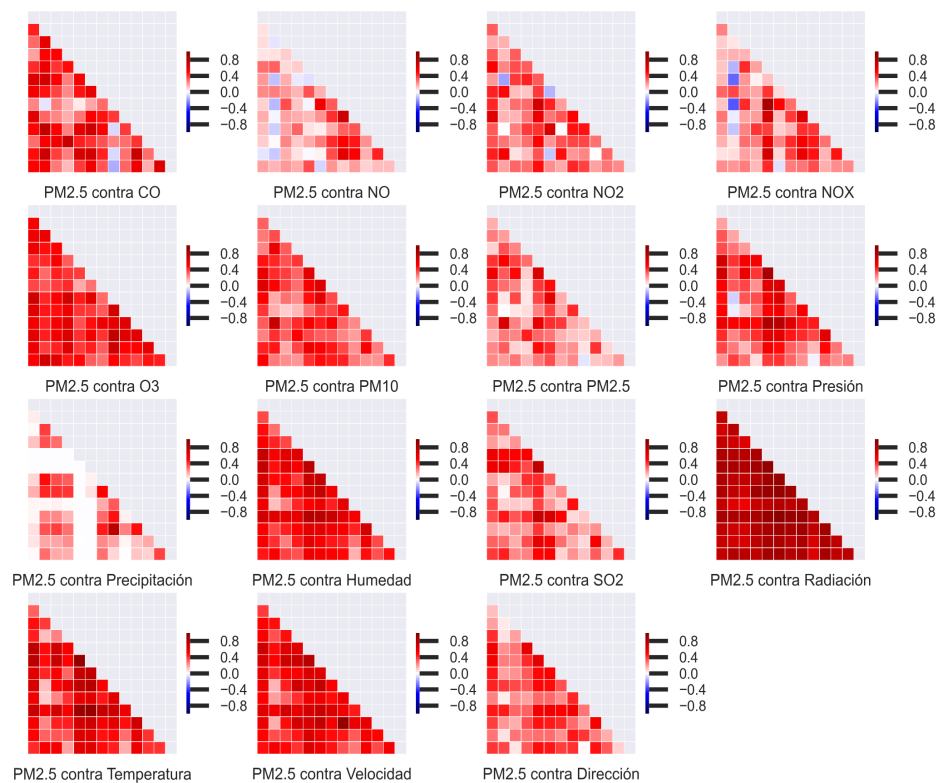
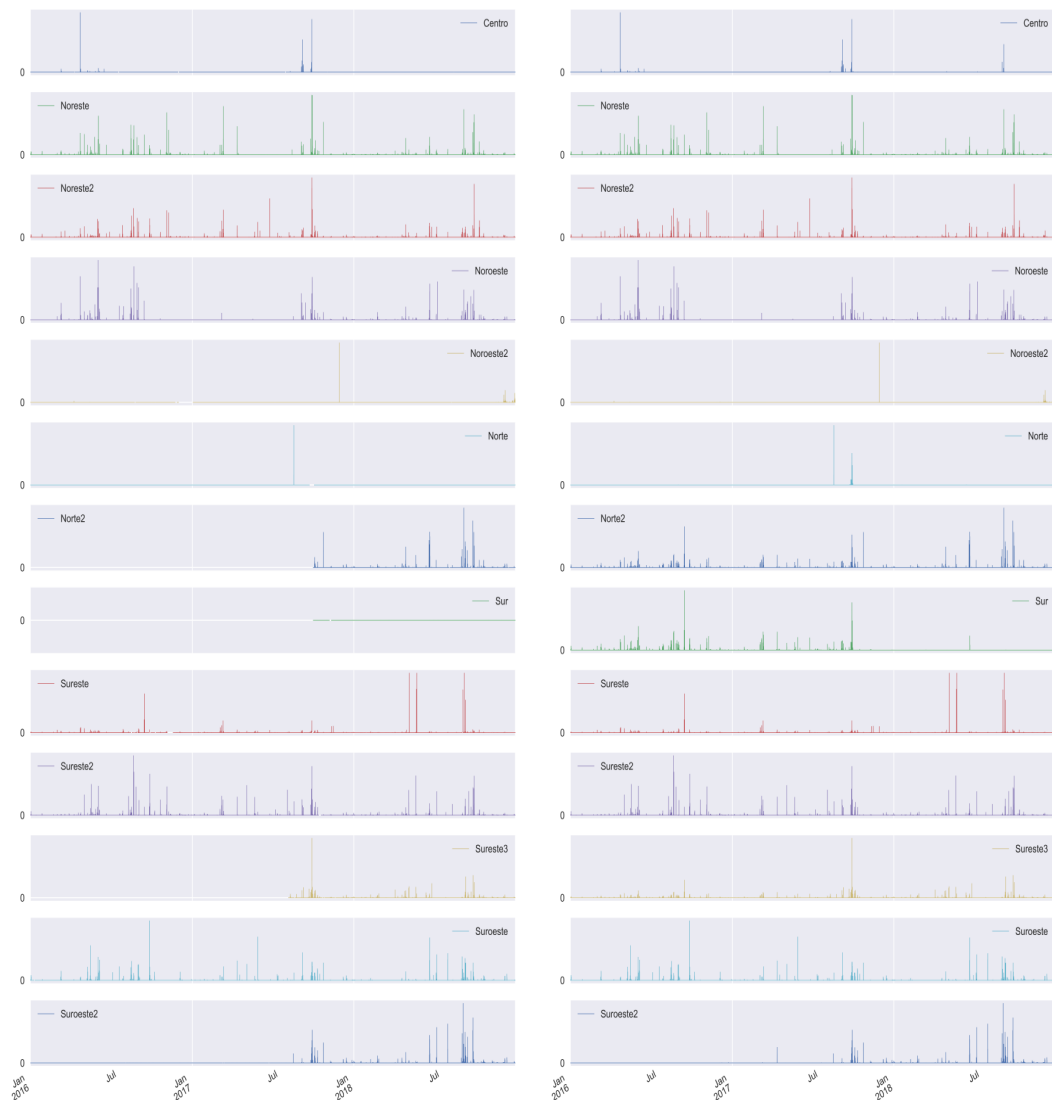


FIGURA 5.33: Matriz de correlación de *presión barométrica* contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* se presentan en algunas estaciones correlaciones altas y correlaciones cercanas a cero (ver figura 5.33), para el resto de las variables la variable O_3 se correlaciona positivamente alto en todas las estaciones.

5.1.9 PRECIPITACIÓN PLUVIAL

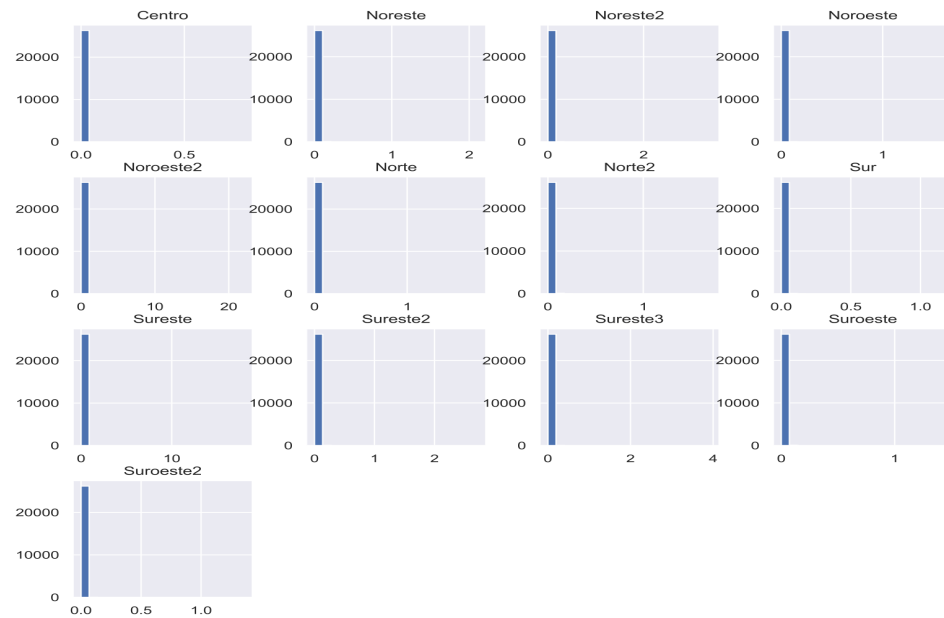


(a) Serie de tiempo de *precipitación pluvial* (b) Serie de tiempo de *precipitación pluvial* relleno de datos desconocidos

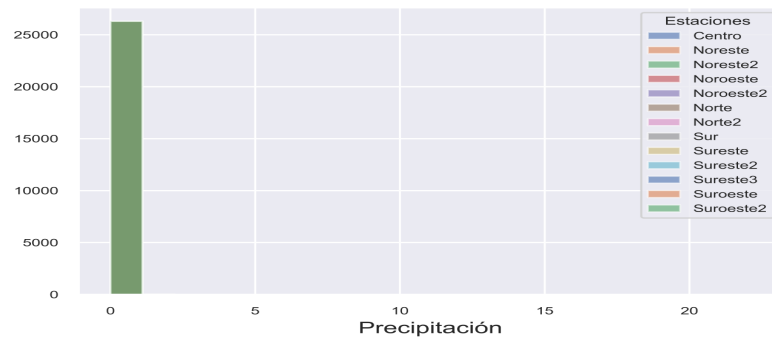
FIGURA 5.34: Series de tiempo de *precipitación pluvial* desde 2016 hasta 2018

La variable *precipitación pluvial* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.34 que la serie de tiempo

es estacionaria ya que parece variar en torno a una media fija.



(a) Histograma de *precipitación pluvial* por estación



(b) Histograma de *precipitación pluvial*

FIGURA 5.35: Histogramas de *precipitación pluvial*

La figura 5.30 muestra los gráficos de series de tiempo para la variable *precipitación pluvial* por estación y la figura 5.35 muestra un histograma de los datos por estación de la variable *precipitación pluvial*. De la figura 5.30 se puede

observar que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.35 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

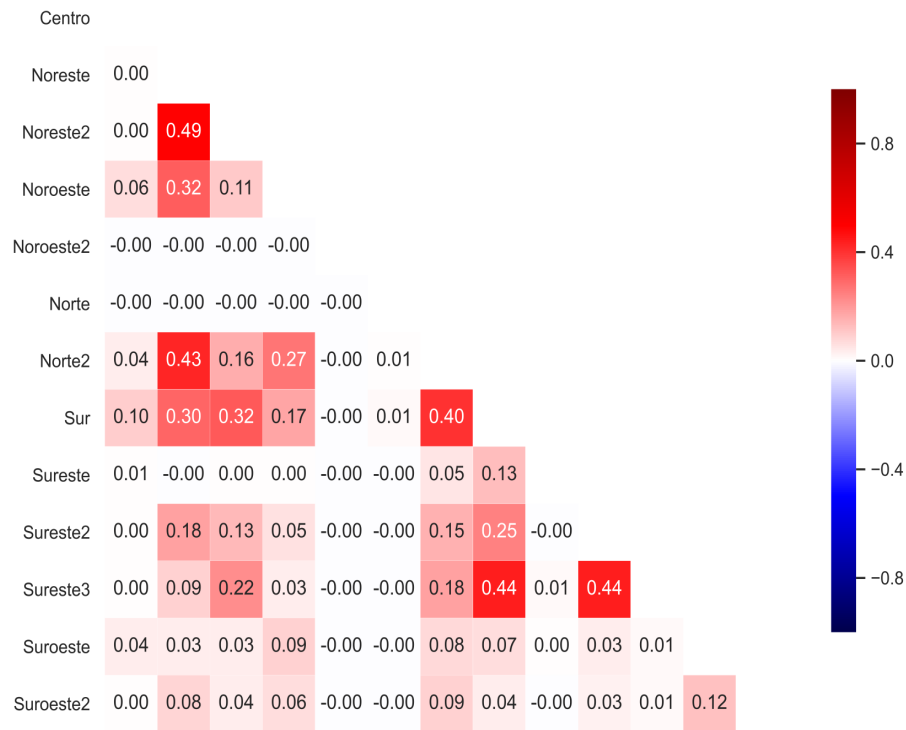


FIGURA 5.36: Matriz de correlación de *precipitación pluvial* entre estaciones

La figura 5.36 muestra que la variable *precipitación pluvial* no se correlaciona con ninguna de las variables restantes.

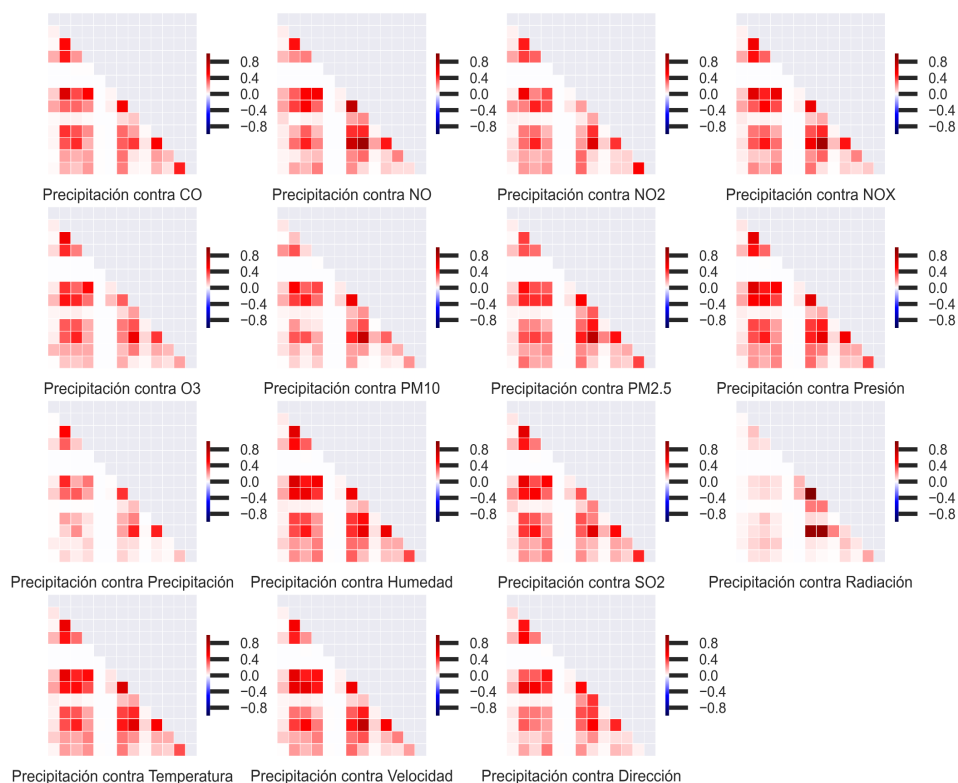
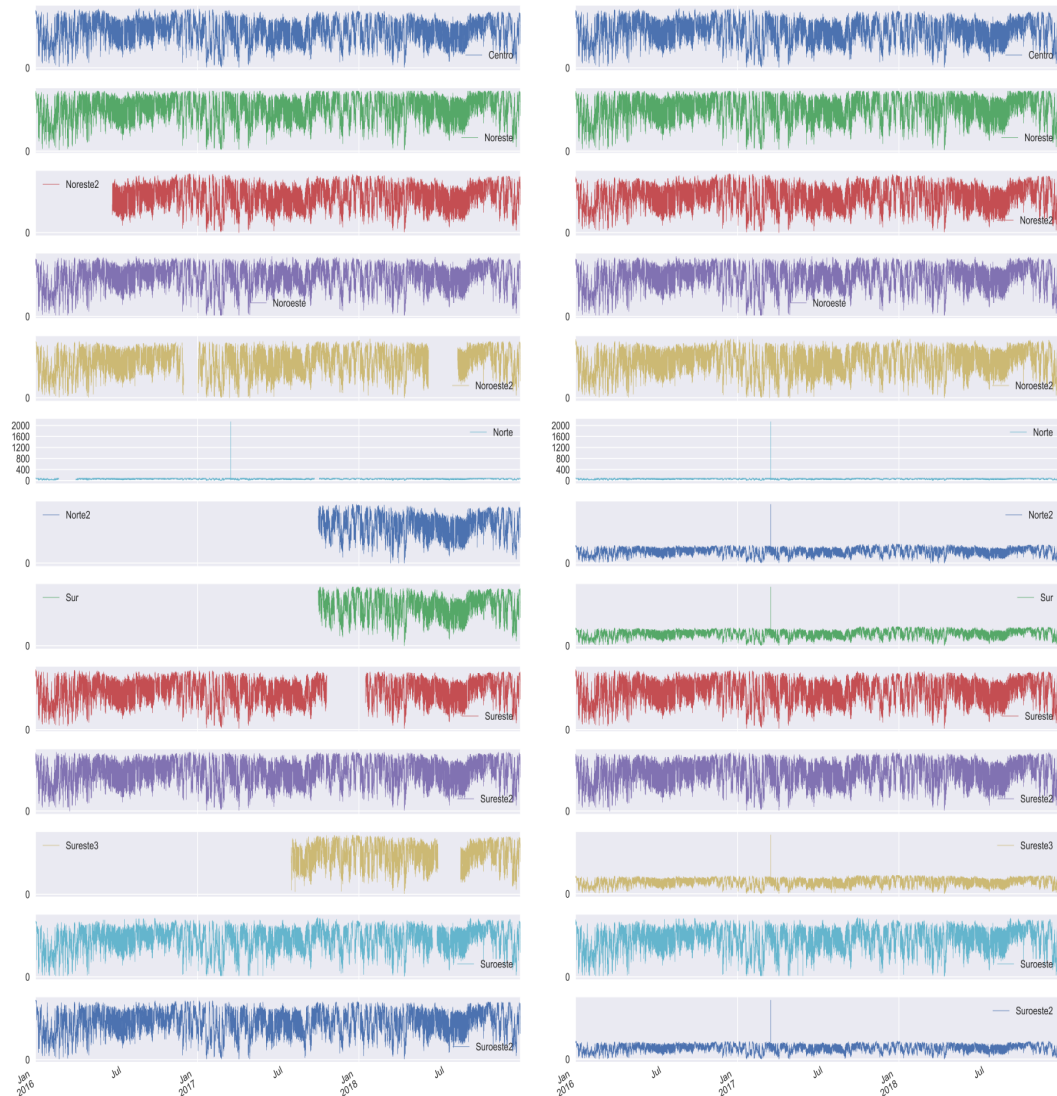


FIGURA 5.37: Matriz de correlación de *precipitación pluvial* contra el resto de las variables

Para las correlaciones de la variable *precipitación pluvial* con el resto de las variables presentan en algunas estaciones correlaciones altas y correlaciones cercanas a cero (ver figura 5.37), es decir, aunque esta variable no se correlaciona con el resto de las variables, en las pocas en las que si lo hace, sucede de manera positiva.

5.1.10 HUMEDAD RELATIVA

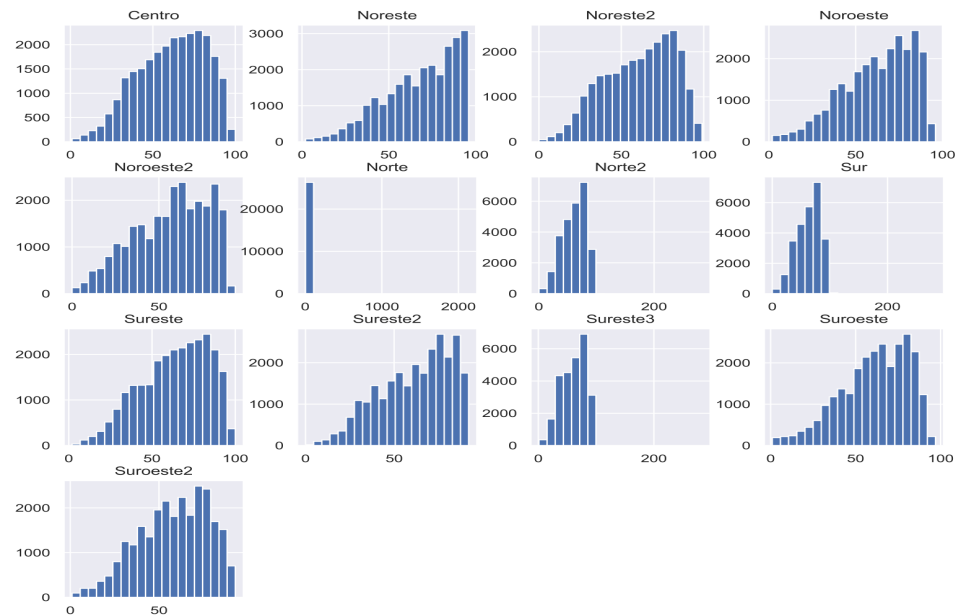


(a) Serie de tiempo de *humedad relativa* original
 (b) Serie de tiempo de *humedad relativa* rellenando datos desconocidos

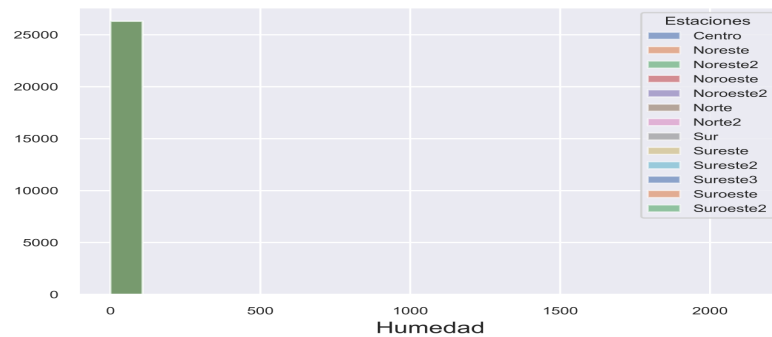
FIGURA 5.38: Series de tiempo de *humedad relativa* desde 2016 hasta 2018

La variable *humedad relativa* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.38 que la serie de

tiempo es estacionaria ya que parece variar entorno a un valor fijo.



(a) Histograma de *humedad relativa* por estación



(b) Histograma de *humedad relativa*

FIGURA 5.39: Histogramas de *humedad relativa*

La figura 5.38 muestra los gráficos de series de tiempo para la variable *humedad relativa* por estación y la figura 5.39 muestra un histograma de los datos por estación de la variable *humedad relativa*. De la figura 5.38 se puede observar

que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.39 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

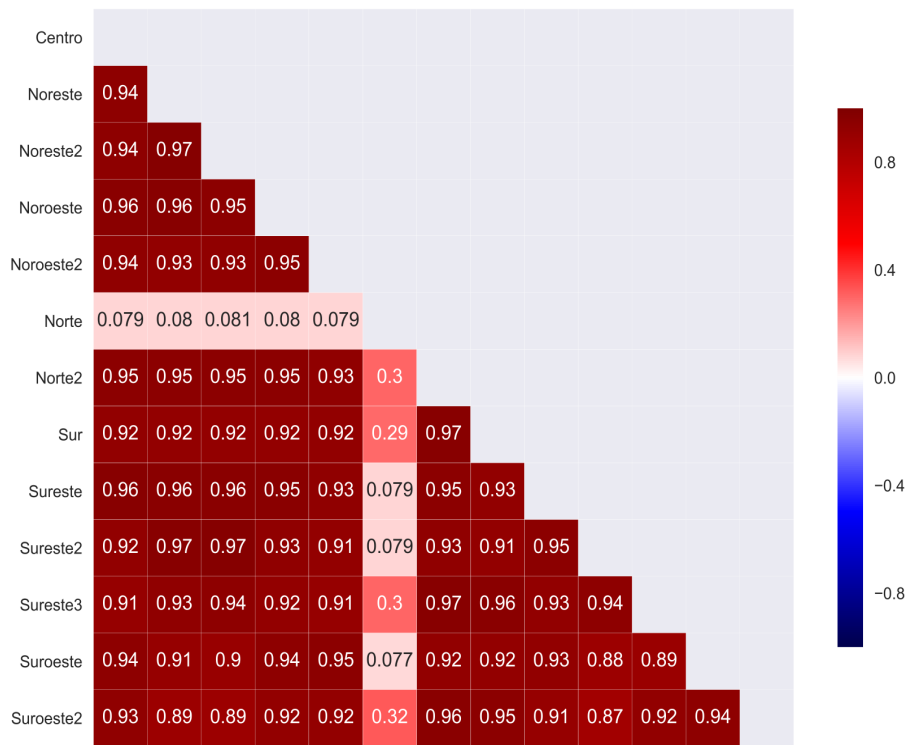


FIGURA 5.40: Matriz de correlación de *humedad relativa* entre estaciones

La figura 5.40 muestra que la variable *humedad relativa* se correlaciona positivamente alto entre la mayoría de las estaciones ya que la estación Norte presenta dos correlaciones positivas bajas contra el resto de las demás estaciones.

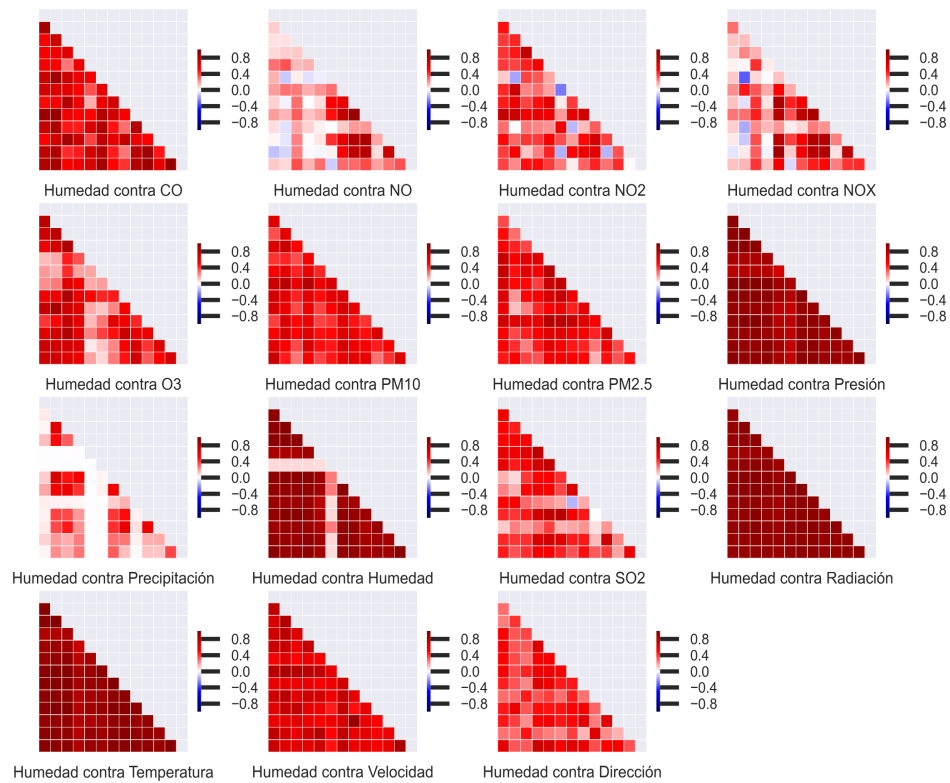
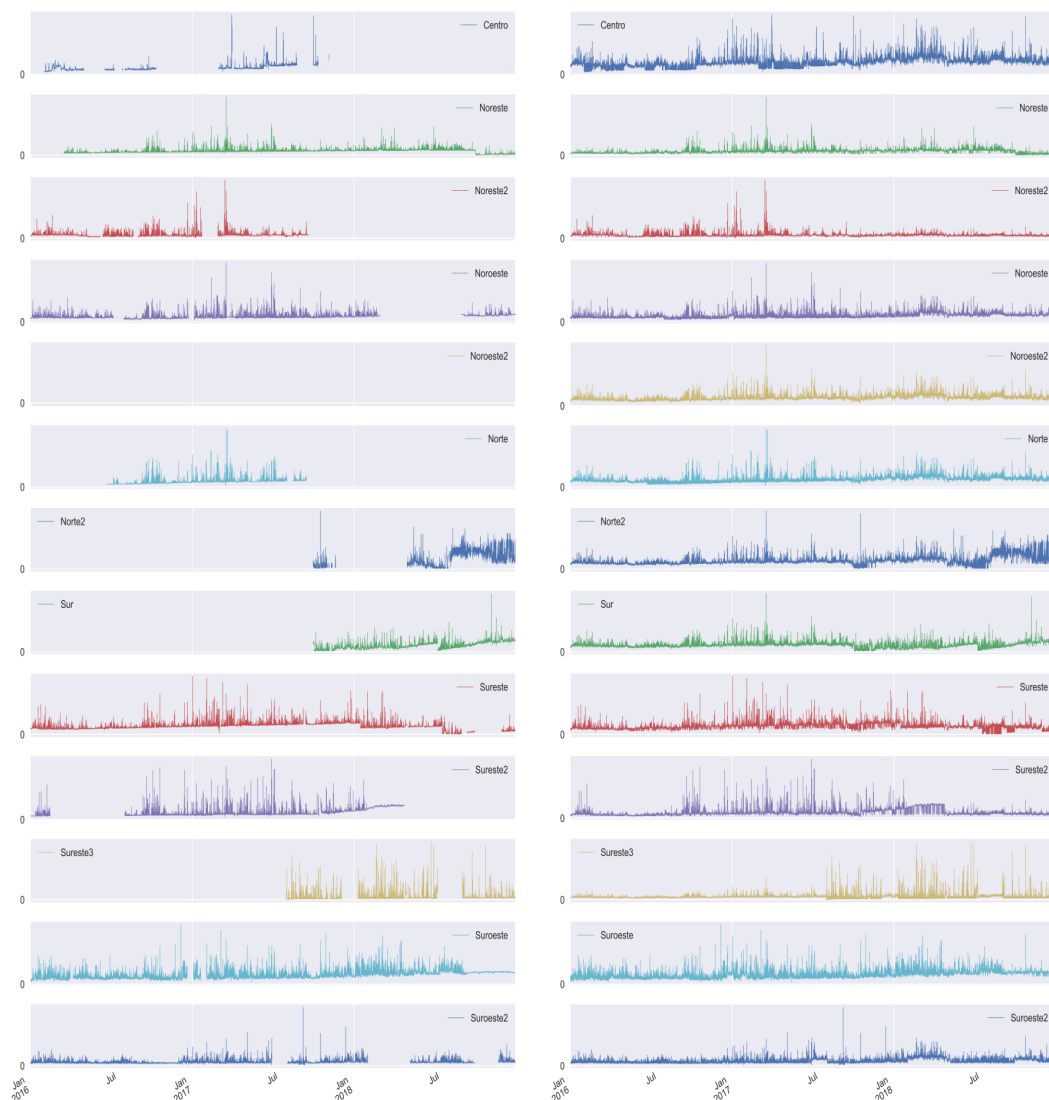


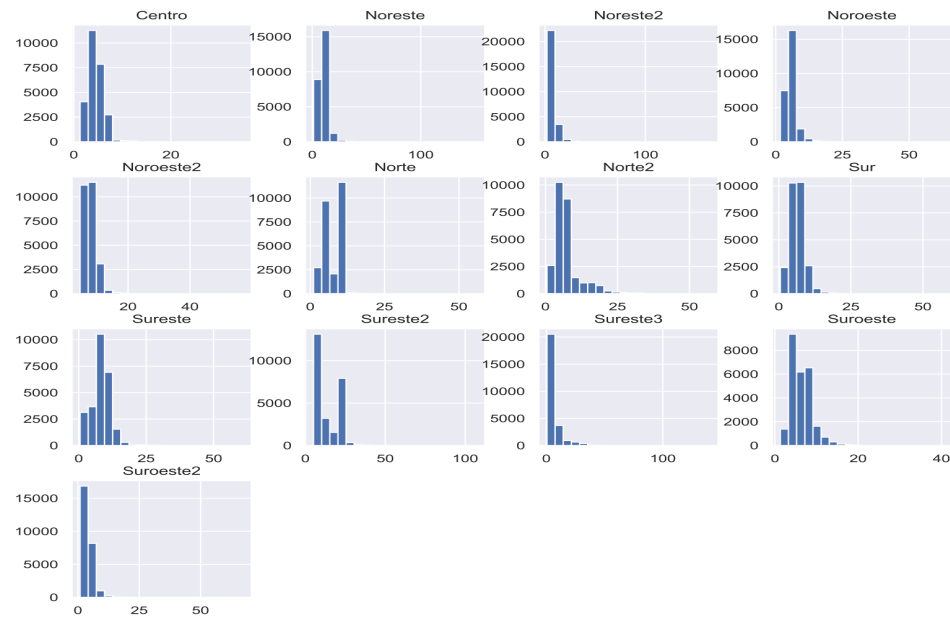
FIGURA 5.41: Matriz de correlación de *humedad relativa* contra el resto de las variables

La variable *humedad relativa* se correlaciona con el resto de las variables positivamente alto en todas las variables restantes (ver figura 5.41).

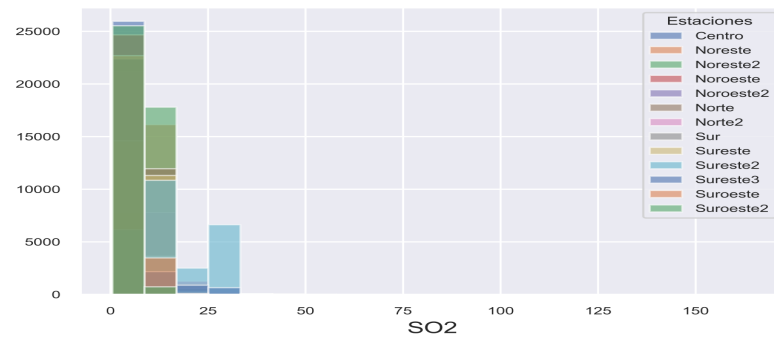
5.1.11 BIÓXIDO DE AZUFRE (SO_2)(a) Serie de tiempo de SO_2 original(b) Serie de tiempo de SO_2 rellenando datos desconocidosFIGURA 5.42: Series de tiempo de SO_2 desde 2016 hasta 2018

La variable SO_2 además de tener datos faltantes que se completan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.42 que la serie de tiempo es

estacionaria ya que parece variar en torno a un valor.



(a) Histograma de SO_2 por estación



(b) Histograma de SO_2

FIGURA 5.43: Histogramas de SO_2

La figura 5.42 muestra los gráficos de series de tiempo para la variable SO_2 por estación y la figura 5.43 muestra un histograma de los datos por estación de la variable SO_2 . De la figura 5.42 se puede observar que si bien las trece

series de tiempo muestran características muy similares, los histogramas de la figura 5.43 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

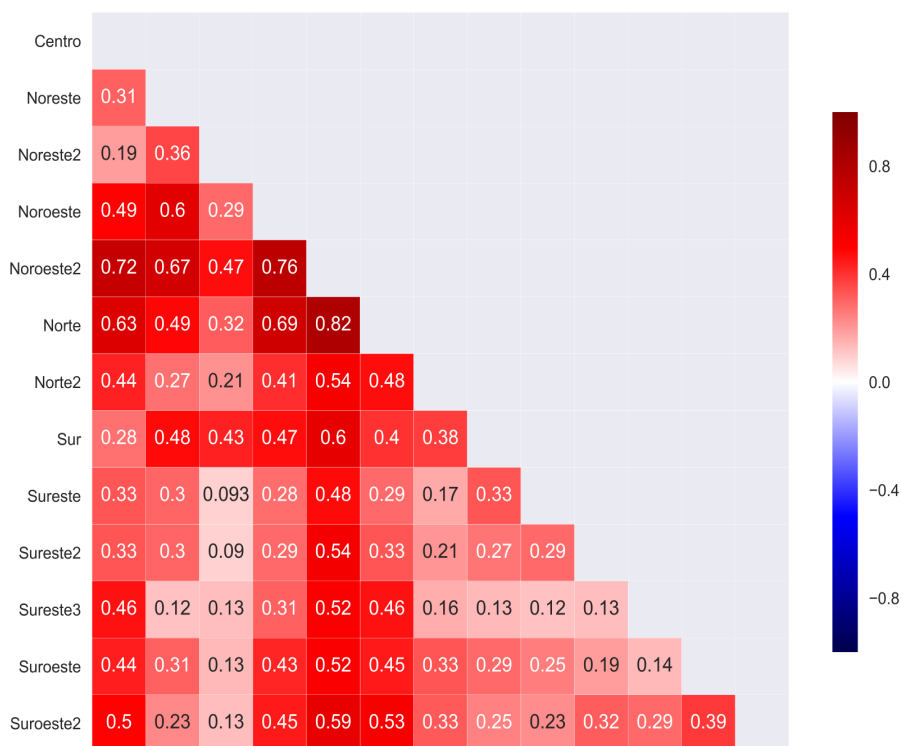


FIGURA 5.44: Matriz de correlación de SO_2 entre estaciones

La figura 5.44 muestra que la variable SO_2 casi no se correlaciona entre estaciones ya que la mayoría de las correlaciones entre estaciones son muy cercanas a cero, mientras que el resto de las estaciones se correlacionan de manera positiva.

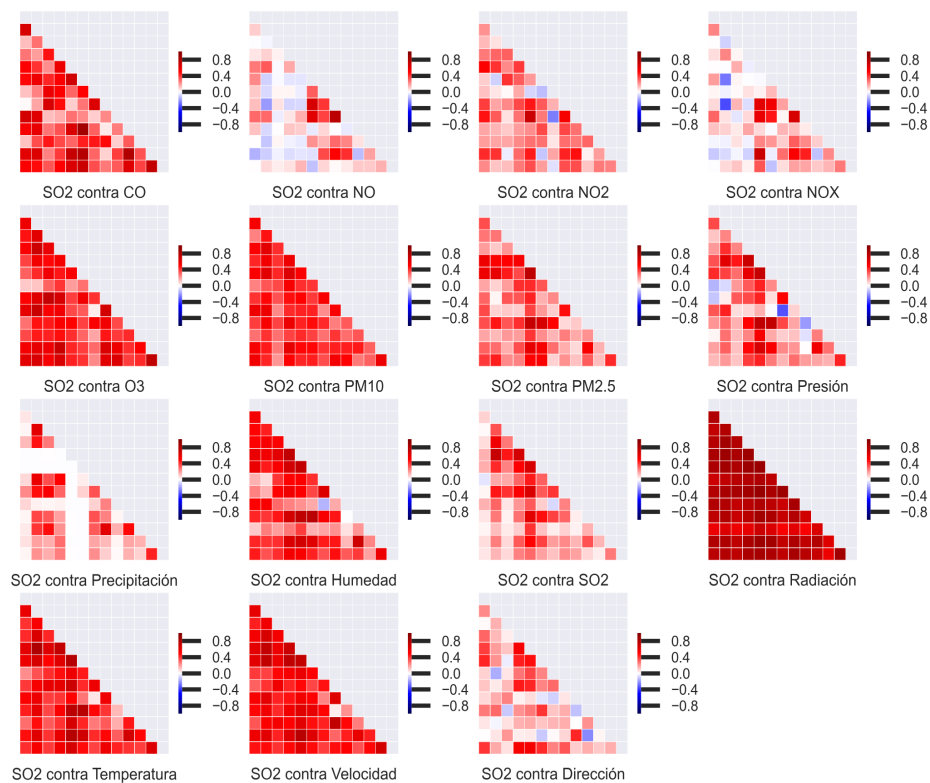
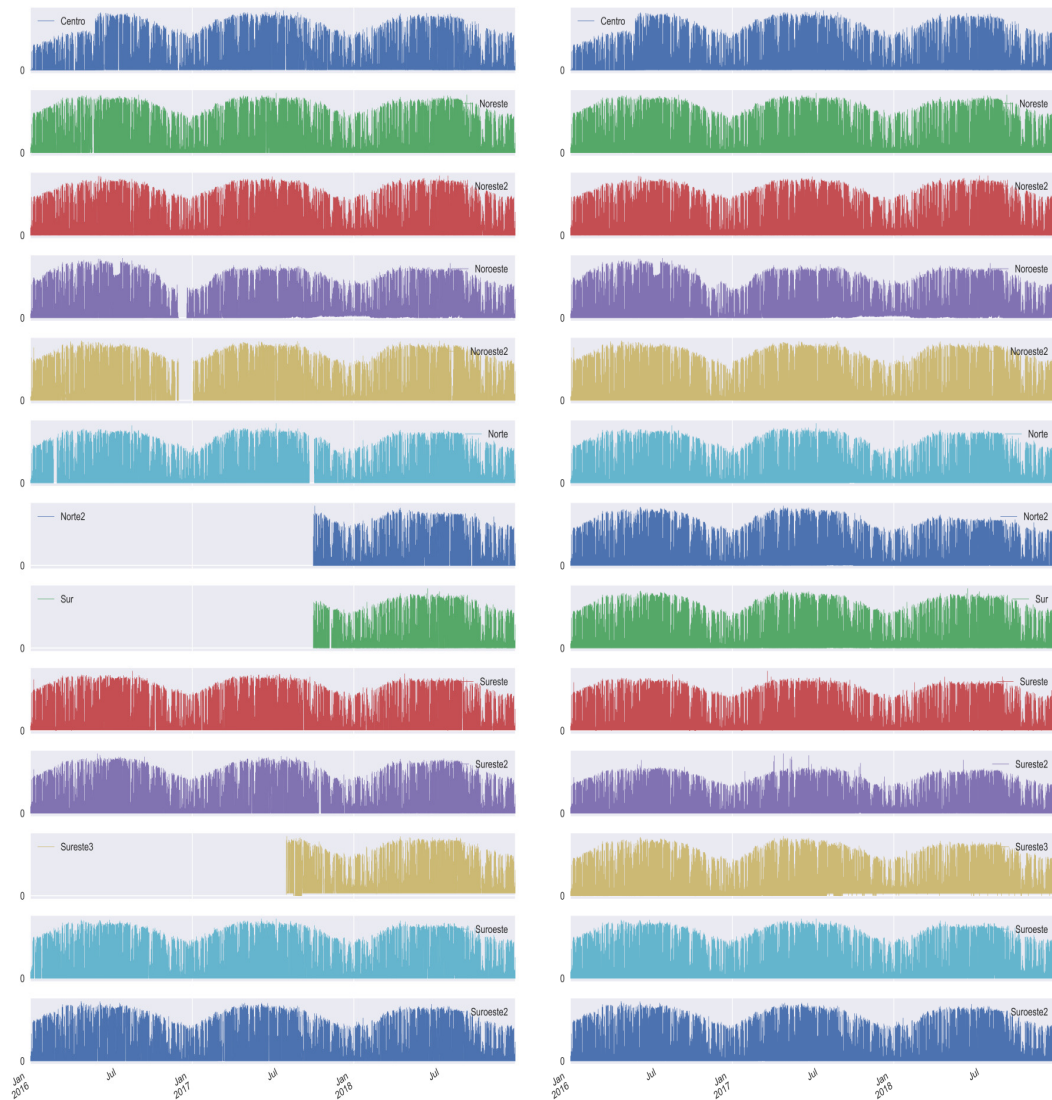


FIGURA 5.45: Matriz de correlación SO_2 contra el resto de las variables

Para las correlaciones de la variable *dirección del viento* se presentan en algunas estaciones correlaciones positivas pequeñas (ver figura 5.45). Para el resto de las variables, la variable SO_2 se correlaciona positivamente alto en todas las estaciones.

5.1.12 RADIACIÓN SOLAR

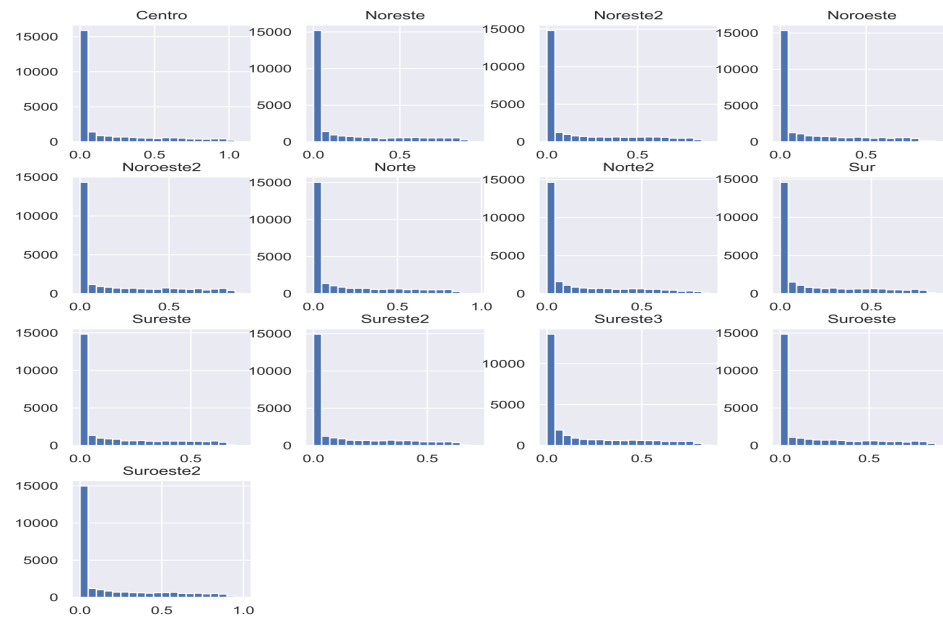


(a) Serie de tiempo de *radiación solar* original (b) Serie de tiempo de *radiación solar* rellenando datos desconocidos

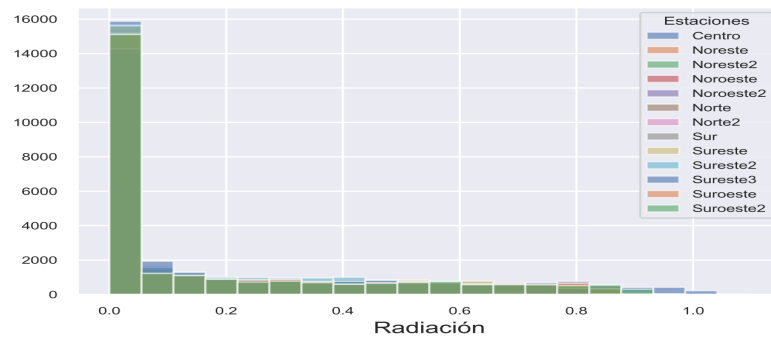
FIGURA 5.46: Series de tiempo de *radiación solar* desde 2016 hasta 2018

La variable *radiación solar* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.46 que la serie de

tiempo es estacionaria ya que parece variar en torno a una media fija.



(a) Histograma de *radiación solar por estación*



(b) Histograma de *radiación solar*

FIGURA 5.47: Histogramas de *radiación solar*

La figura 5.46 muestra los gráficos de series de tiempo para la variable *radiación solar* por estación y la figura 5.47 muestra un histograma de los datos por estación de la variable *radiación solar*. De la figura 5.46 se puede observar

que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.47 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

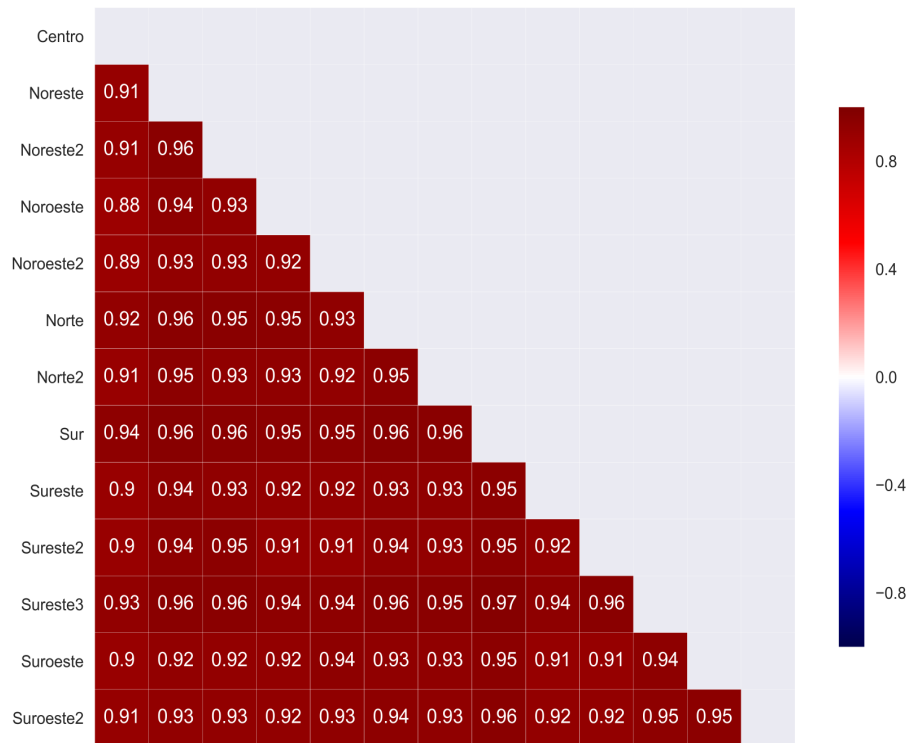


FIGURA 5.48: Matriz de correlación de *radiación solar* entre estaciones

La figura 5.48 muestra que la variable *radiación solar* se correlaciona positivamente alto entre todas las estaciones.

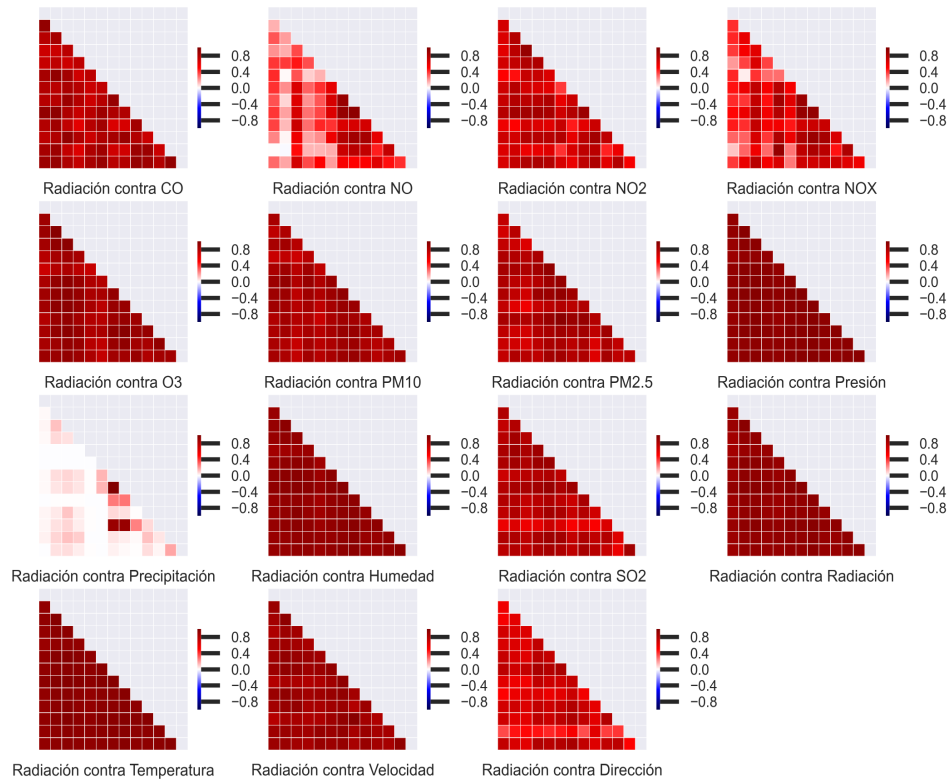
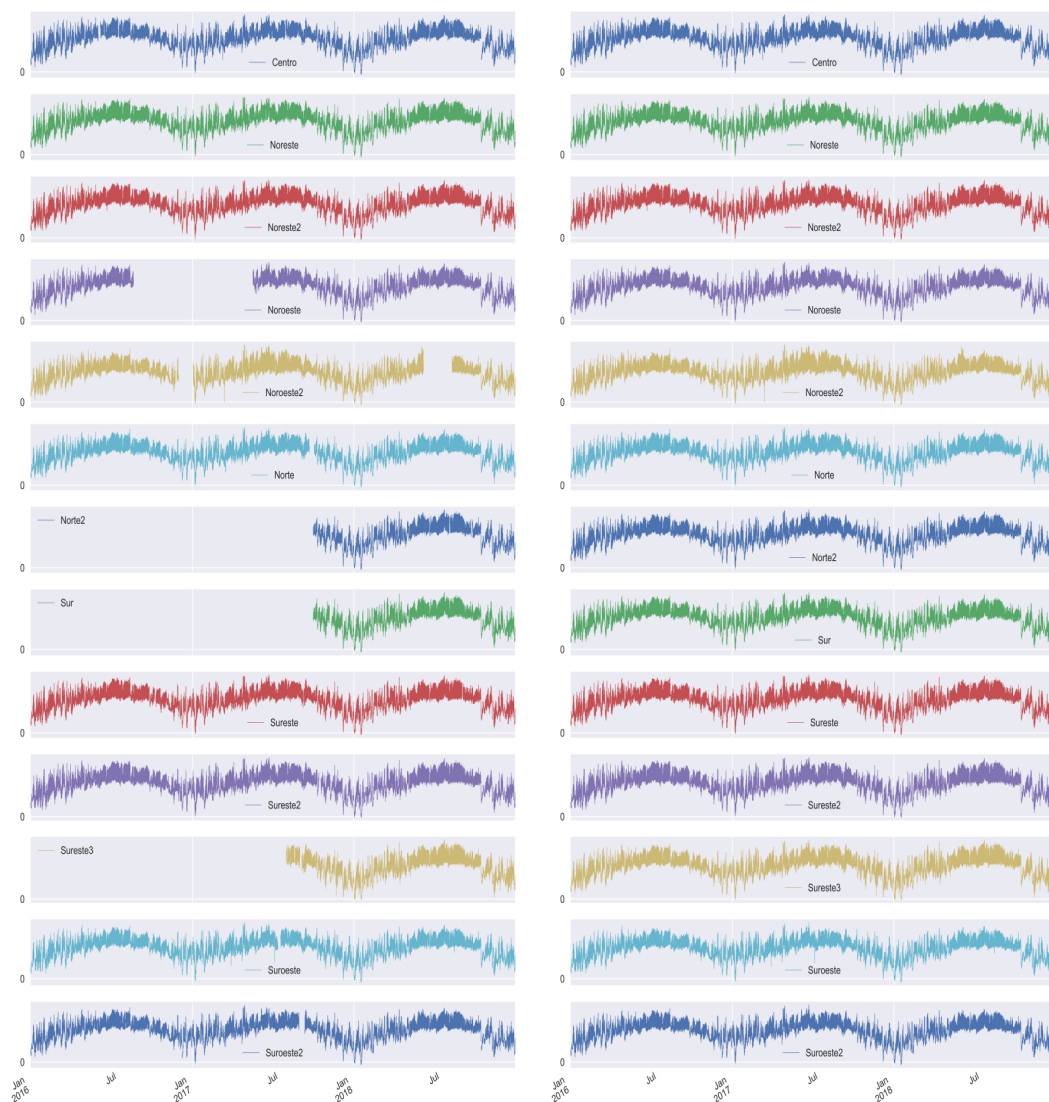


FIGURA 5.49: Matriz de correlación de *radiación solar* contra el resto de las variables

Para las correlaciones de la variable *radiación solar* se presentan correlaciones positivas altas para todas las estaciones en todas cuando se hace contra el resto de variables restantes (ver figura 5.49).

5.1.13 TEMPERATURA AMBIENTAL

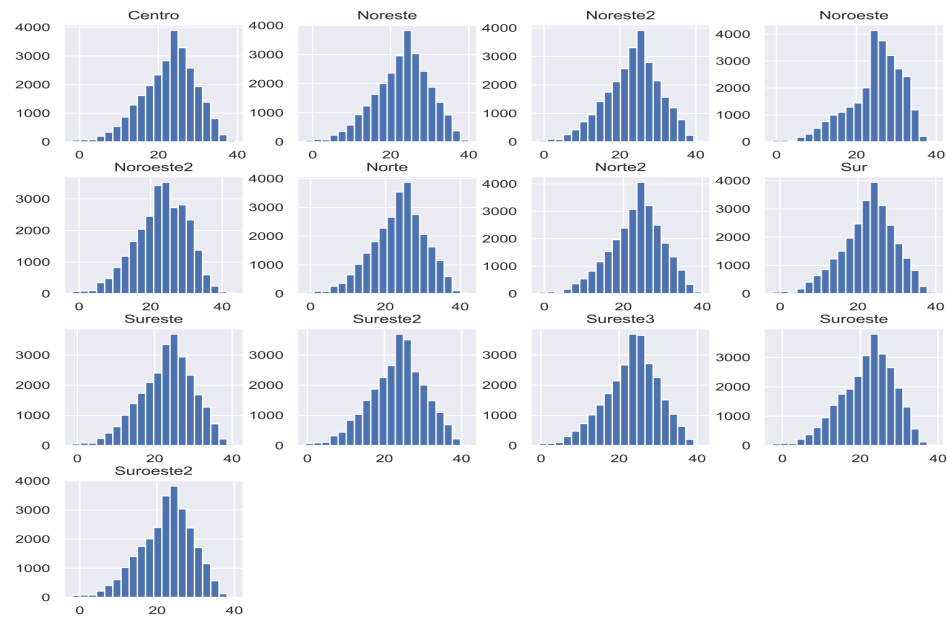


(a) Serie de tiempo de *temperatura ambiental* original
 (b) Serie de tiempo de *temperatura ambiental* rellenando datos desconocidos

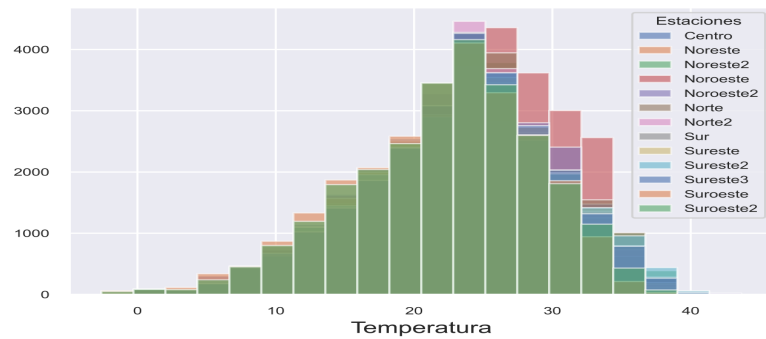
FIGURA 5.50: Series de tiempo de *temperatura ambiental* desde 2016 hasta 2018

La variable *temperatura ambiental* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.50 que la serie de

tiempo es estacionaria ya que parece variar en torno a una media fija.



(a) Histograma de *temperatura ambiental* por estación



(b) Histograma de *temperatura ambiental*

FIGURA 5.51: Histogramas de *temperatura ambiental*

La figura 5.50 muestra los gráficos de series de tiempo para la variable *temperatura ambiental* por estación y la figura 5.51 muestra un histograma de los datos por estación de la variable *temperatura ambiental*. De la figura 5.50 se puede

observar que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.51 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

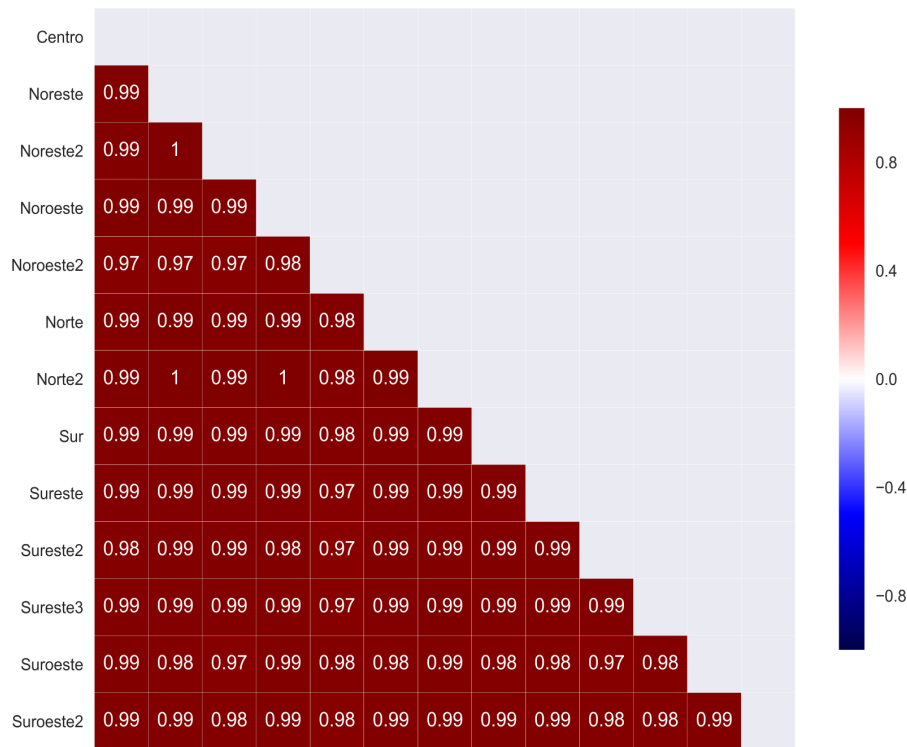


FIGURA 5.52: Matriz de correlación de *temperatura ambiental* entre estaciones

La figura 5.52 muestra que la variable *temperatura ambiental* se correlaciona positivamente alto entre todas las estaciones.

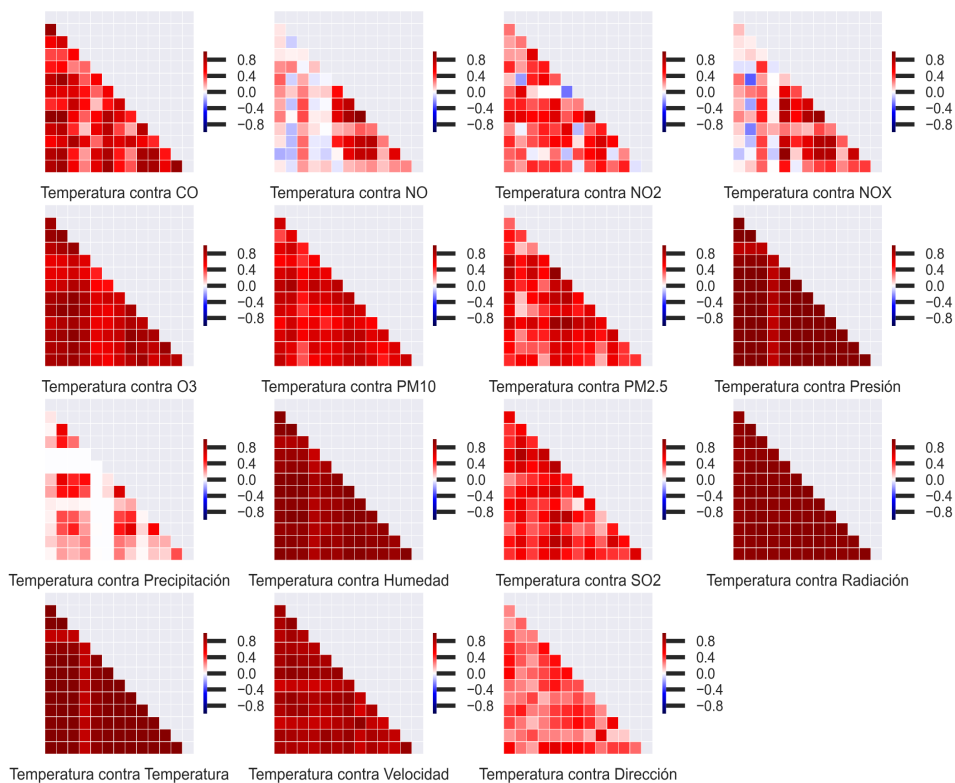
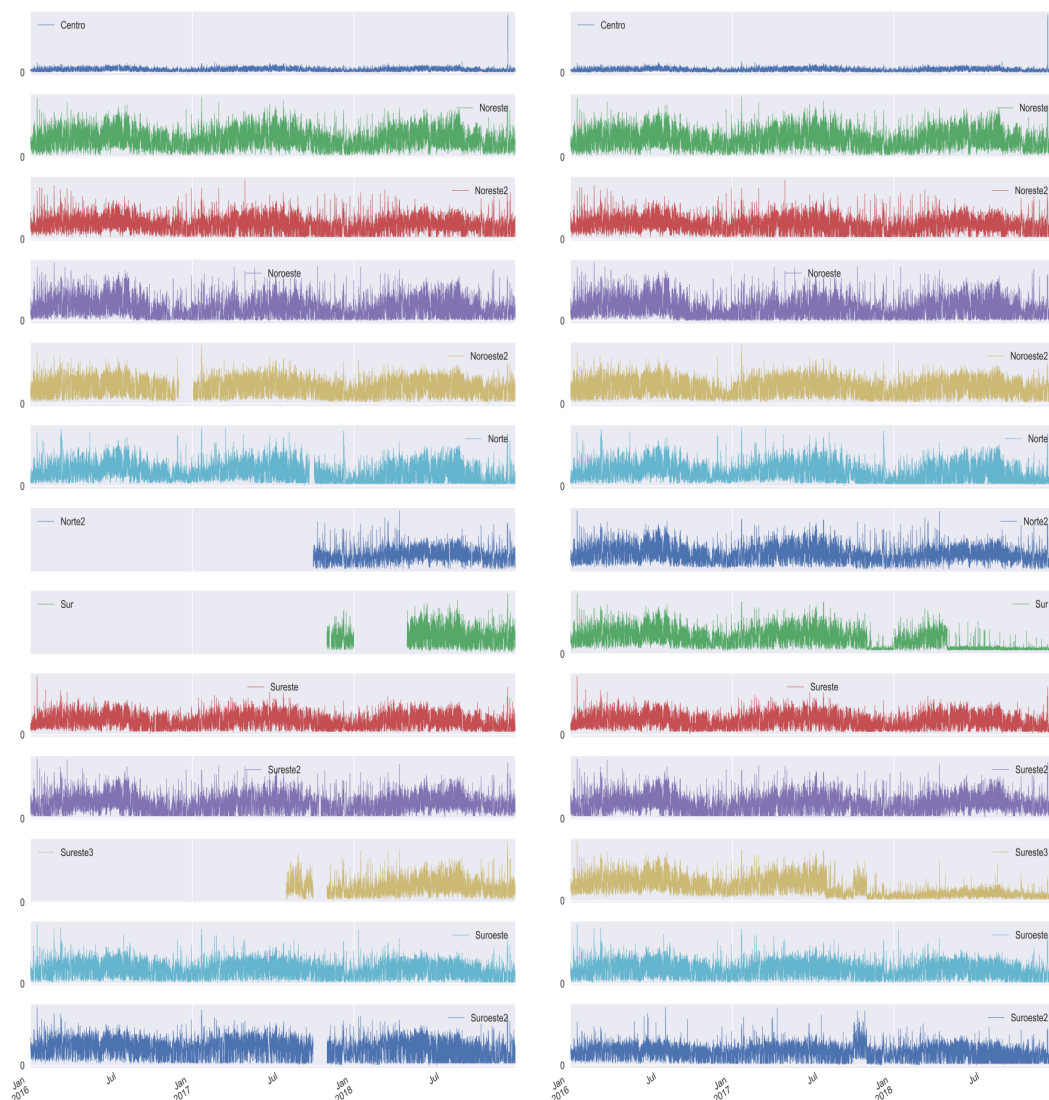


FIGURA 5.53: Matriz de correlación de *temperatura ambiental* contra el resto de las variables

Para las correlaciones de la variable *temperatura ambiental* se presentan correlaciones positivas altas para todas las estaciones en todas cuando se hace contra el resto de variables restantes (ver figura 5.53).

5.1.14 VELOCIDAD DEL VIENTO

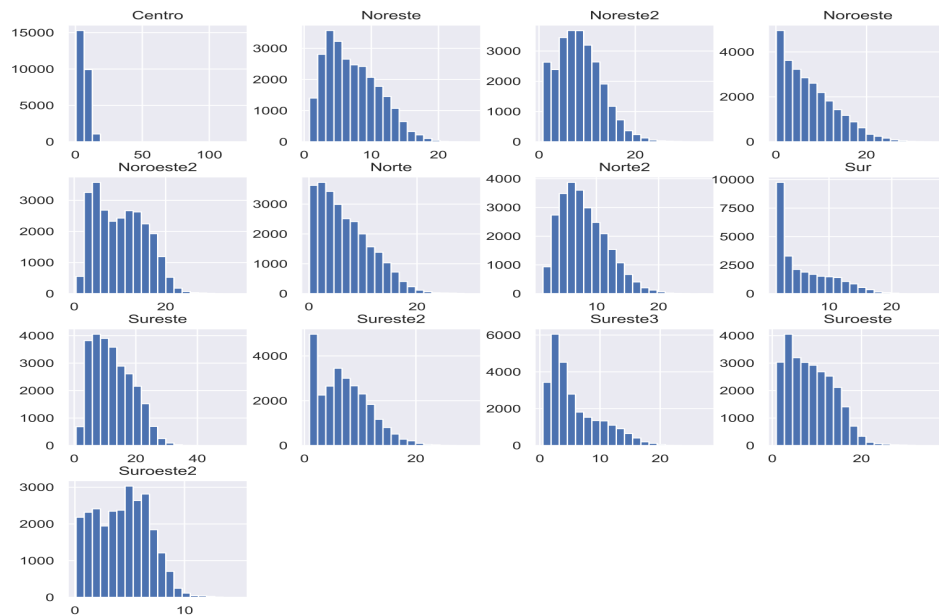


(a) Serie de tiempo de *velocidad del viento* original
 (b) Serie de tiempo de *velocidad del viento* relleno de datos desconocidos

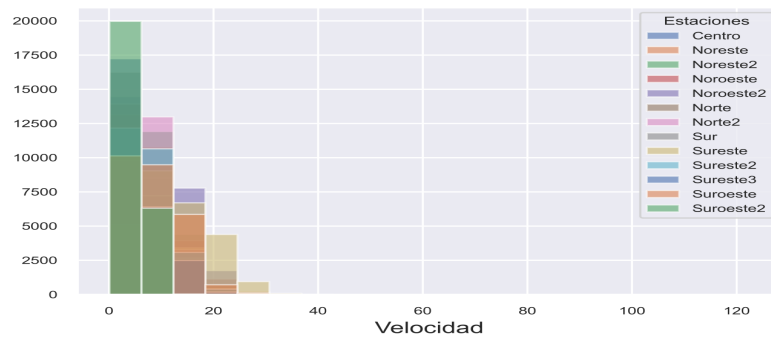
FIGURA 5.54: Series de tiempo de *velocidad del viento* desde 2016 hasta 2018

La variable *velocidad del viento* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.54 que la serie de tiempo

es estacionaria ya que parece variar en torno a un valor constante.



(a) Histograma de *velocidad del viento* por estación



(b) Histograma de *velocidad del viento*

FIGURA 5.55: Histogramas de *velocidad del viento*

La figura 5.54 muestra los gráficos de series de tiempo para la variable *velocidad del viento* por estación y la figura 5.55 muestra un histograma de los datos por estación de la variable *velocidad del viento*. De la figura 5.54 se puede observar

que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.55 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características principales de los datos que dependen del tiempo.

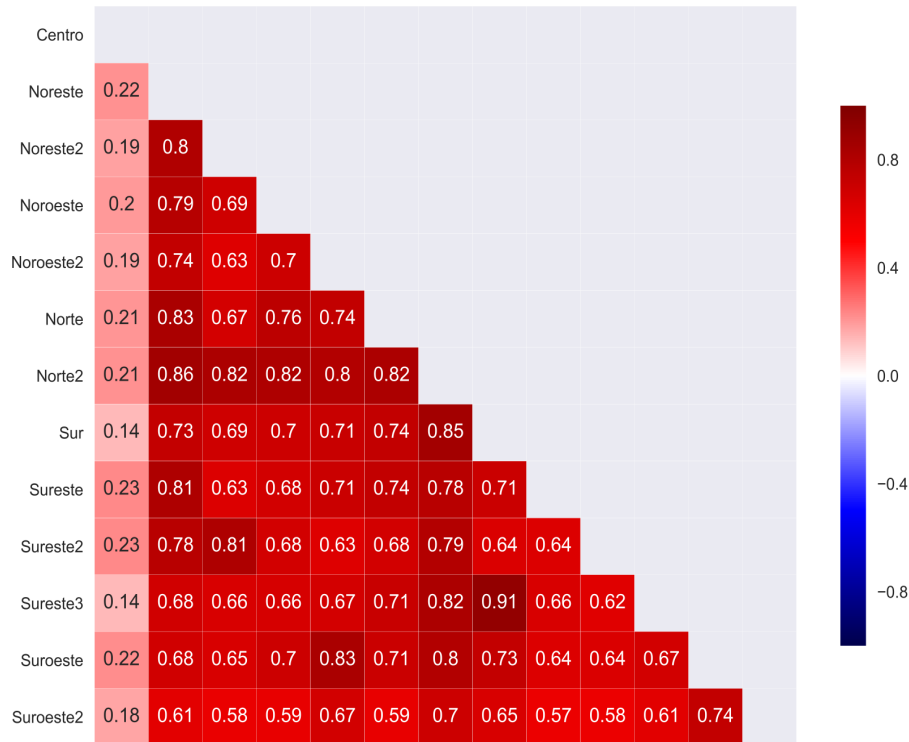


FIGURA 5.56: Matriz de correlación de *velocidad del viento* entre estaciones

La figura 5.56 muestra que la variable *velocidad del viento* se correlaciona positivamente alto entre todas las estaciones, menos en la estación Centro ya que presenta correlaciones menores a 0.2 entre la mayoría de las estaciones

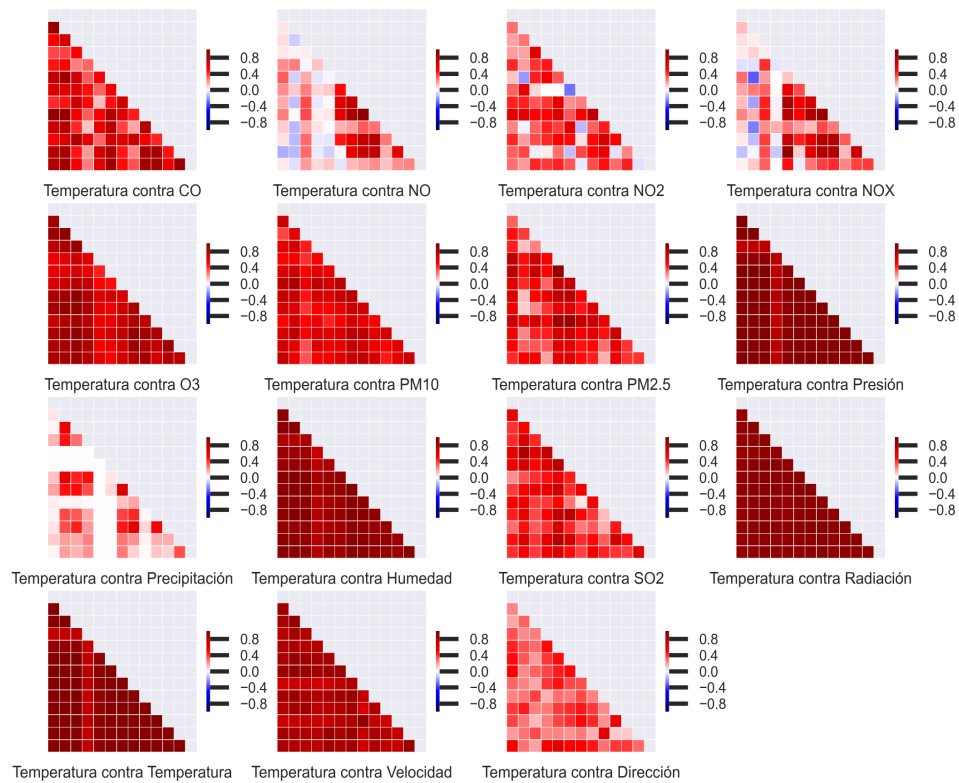
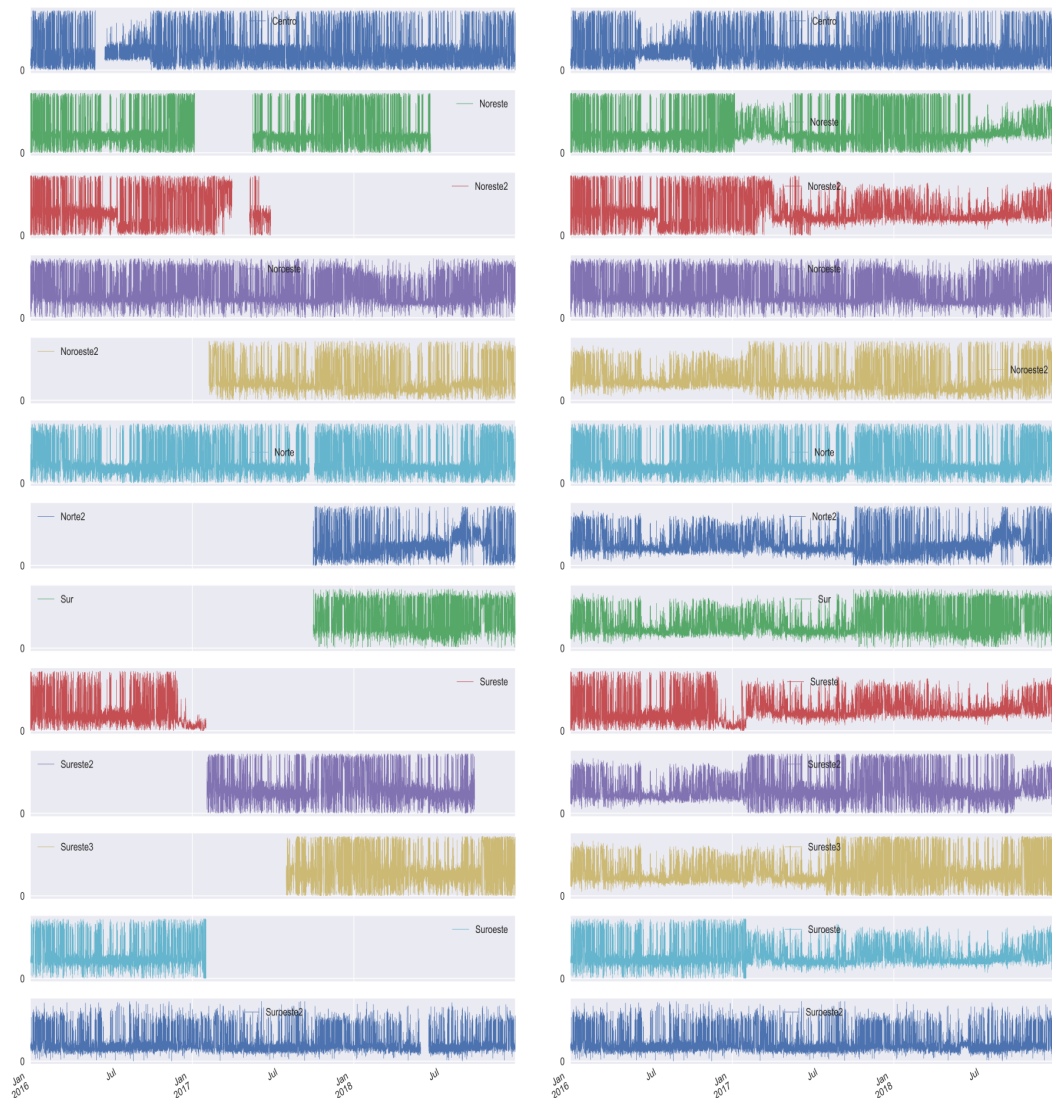


FIGURA 5.57: Matriz de correlación de *velocidad del viento* contra el resto de las variables

Para las correlaciones de la variable *velocidad del viento* se presentan correlaciones positivas altas para todas las estaciones en todas cuando se hace contra el resto de variables restantes (ver figura 5.57).

5.1.15 DIRECCIÓN DEL VIENTO

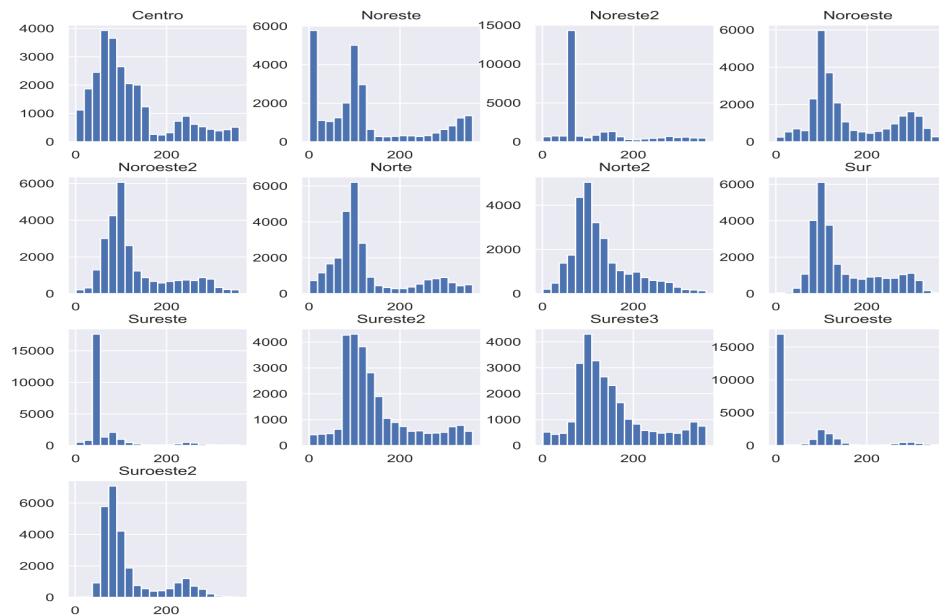


(a) Serie de tiempo de *dirección del viento* original
 (b) Serie de tiempo de *dirección del viento* relleno llenando datos desconocidos

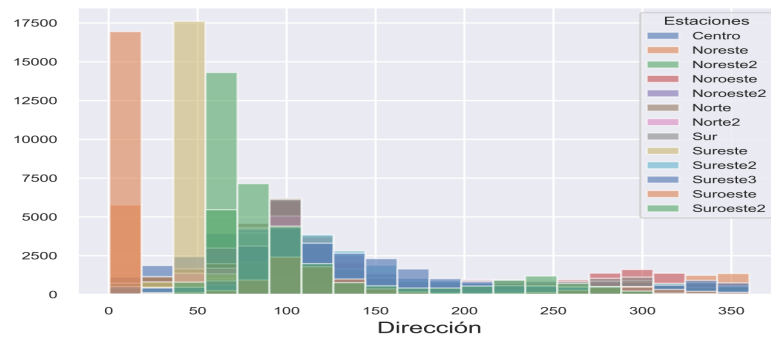
FIGURA 5.58: Series de tiempo de *dirección del viento* desde 2016 hasta 2018

La variable *dirección del viento* además de tener datos faltantes que se rellenan haciendo el promedio de las estaciones cercanas y asignando el promedio a las estaciones sin registros, se puede apreciar en la figura 5.58 que la serie de

tiempo es estacionaria ya que parece variar en torno a una media.



(a) Histograma de *dirección del viento* por estación



(b) Histograma de *dirección del viento*

FIGURA 5.59: Histogramas de *dirección del viento*

La figura 5.58 muestra los gráficos de series de tiempo para la variable *dirección del viento* por estación y la figura 5.59 muestra un histograma de los datos por estación de la variable *dirección del viento*. De la figura 5.58 se puede observar

que si bien las trece series de tiempo muestran características muy similares, los histogramas de la figura 5.59 son diferentes, ya que el histograma resume los datos a través de la dimensión del tiempo, y al hacerlo, se pierden las características clave de los datos que dependen del tiempo.

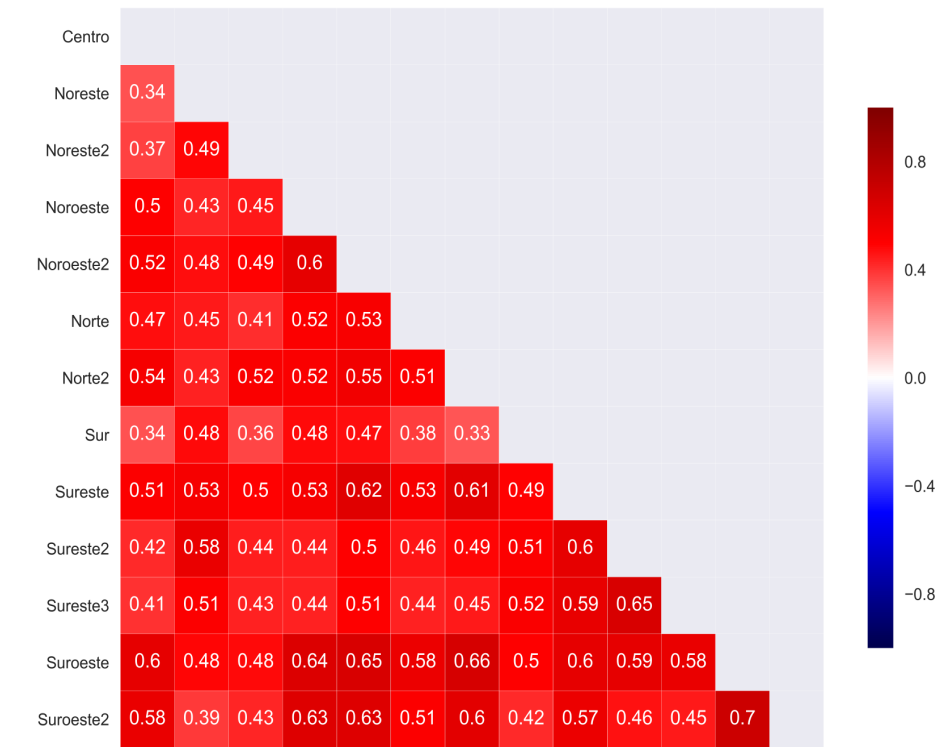


FIGURA 5.60: Matriz de correlación de *dirección del viento* entre estaciones

Para las correlaciones de la variable *dirección del viento* se presentan correlaciones positivas altas (ver figura 5.61).

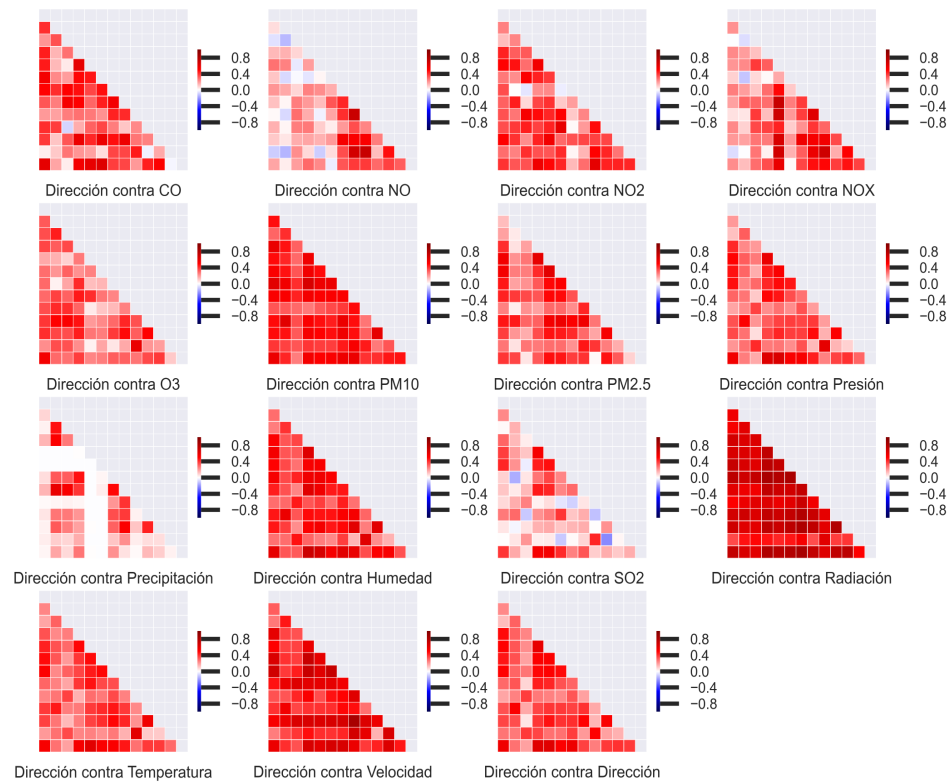


FIGURA 5.61: Matriz de correlación de *dirección del viento* contra el resto de las variables

5.2. Medición del Error

A partir de una *validación cruzada* (eliminar una o varias estaciones para interpolar su valor con base en la información restante) y se comparan los valores reales e interpolados a partir de los métodos:

- Error porcentual absoluto medio

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \frac{|\hat{y}_{x_i,t} - y_{x_i,t}|}{y_{x_i,t}}.$$

- Error absoluto medio

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{y}_{x_i,t} - y_{x_i,t}|.$$

- Error cuadrático medio

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^T (\hat{y}_{x_i,t} - y_{x_i,t})^2.$$

- Error cuadrático medio

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_{x_i,t} - y_{x_i,t})^2},$$

donde $y_{x_i,t}$ es el valor real en la estación x_i al tiempo t , $\hat{y}_{x_i,t}$ es el valor predicho en la estación x_i al tiempo t y T es el número total de iteraciones.

5.2.1 ERROR

Se define el *error* entre un pronóstico contra el valor real observado como:

$$e_t = \hat{y}_{x_i,t} - y_{x_i,t},$$

donde $y_{x_i,t}$ es el valor real en la estación x_i al tiempo t , $\hat{y}_{x_i,t}$ es el valor predicho en la estación x_i al tiempo t .

Con esta definición, si el pronóstico supera el valor real, el error será positivo y si el pronóstico no supera la demanda el error será negativo.

5.2.2 MAPE

El *error porcentual absoluto medio* (MAPE) es una medida de error que se encuentra entre los más utilizados para medir la precisión del pronóstico.

La idea de MAPE consiste en la suma de los errores absolutos individuales divididos por el valor real (en cada período por separado) o el promedio de los errores porcentuales. Un problema que se puede ver de esta métrica, es que MAPE divide cada error individual por el valor real, por lo que está sesgado, es decir, los errores altos durante los periodos de valores reales pequeños tendrán un gran impacto en el valor de MAPE, además de presentarse el caso en el que el valor real observado sea cero, el MAPE producirá errores al ser evaluado.

5.2.3 MAE

El *error absoluto medio* (MAE) es uno de las métricas también más utilizadas para medir la precisión del pronóstico. La idea de MAE es hacer la media del error absoluto.

Un problema que presenta esta métrica es que no se ajusta a los valores reales promedio, por ejemplo si al calcular MAE el resultado es diez, no se sabe si esto es bueno o malo, es decir, si sus valores reales promedio es de mil, el MAE es muy bueno, pero si los errores reales promedio son uno, la precisión en este caso es mala. Para tratar con este problema se suele dividir MAE entre los valores reales promedio para obtener un porcentaje de la siguiente manera:

$$\text{MAEP} = \frac{\text{MAE}}{\frac{1}{n} \sum_{t=0}^n y_{x_i,t}},$$

al cual se llamará para este trabajo de tesis *error absoluto medio porcentual* (MAEP).

5.2.4 RMSE

El *error cuadrático medio* (RMSE) es una métrica que se define como la raíz cuadrada del error cuadrado promedio, entonces RMSE al igual que MAE no se adapta al promedio de los valores reales. Para tratar con este problema se suele hacer:

$$\text{RMSEP} = \frac{\text{RMSE}}{\frac{1}{n} \sum_{t=0}^n y_{x_i,t}},$$

al cual se llamará *error cuadrático medio porcentual* (RMSEP). A diferencia del MAE, el RMSE no trata cada error de la misma manera, es decir, da más importancia a los errores más grandes. Eso implica que un gran error basta para obtener un mal RMSE.

5.2.5 MSE

El *error cuadrático medio* (MSE) es utilizado por muchos algoritmos ya que es más rápido y fácil de calcular que otras métricas, pero al estar el error al cuadrado, MSE no se ajusta al valor del error original, lo que resulta en una métrica que no se puede relacionar directamente entre la predicción y el valor real observado.

En conclusión MAE brinda protección contra los valores atípicos, mientras que RMSE brinda la seguridad de obtener un pronóstico imparcial. Si se usa MAE como métrica de error dará como resultado un alto sesgo, por lo que es probable que se escoja usar RMSE. Por otro lado, si el conjunto de datos contiene muchos valores atípicos, resultará en un pronóstico sesgado, por lo que es posible que se utilice la métrica MAE.

Para este trabajo de tesis se evalúan las siguientes métricas mencionadas en el

orden siguiente:

1. MAPE
2. MAE
3. MAEP
4. RMSE
5. RMSEP
6. MSE

5.3. Entorno Computacional y Tiempo de Ejecución

Para la parte de programación de este trabajo de tesis se utiliza para ejecutar el código correspondiente (Serna, 2020), una computadora con sistema operativo Windows 10 Home de 64 bits, procesador Intel Core i-7 2.2 GHz y 8 GB de RAM. Lenguaje de programación Python versión 3.7.

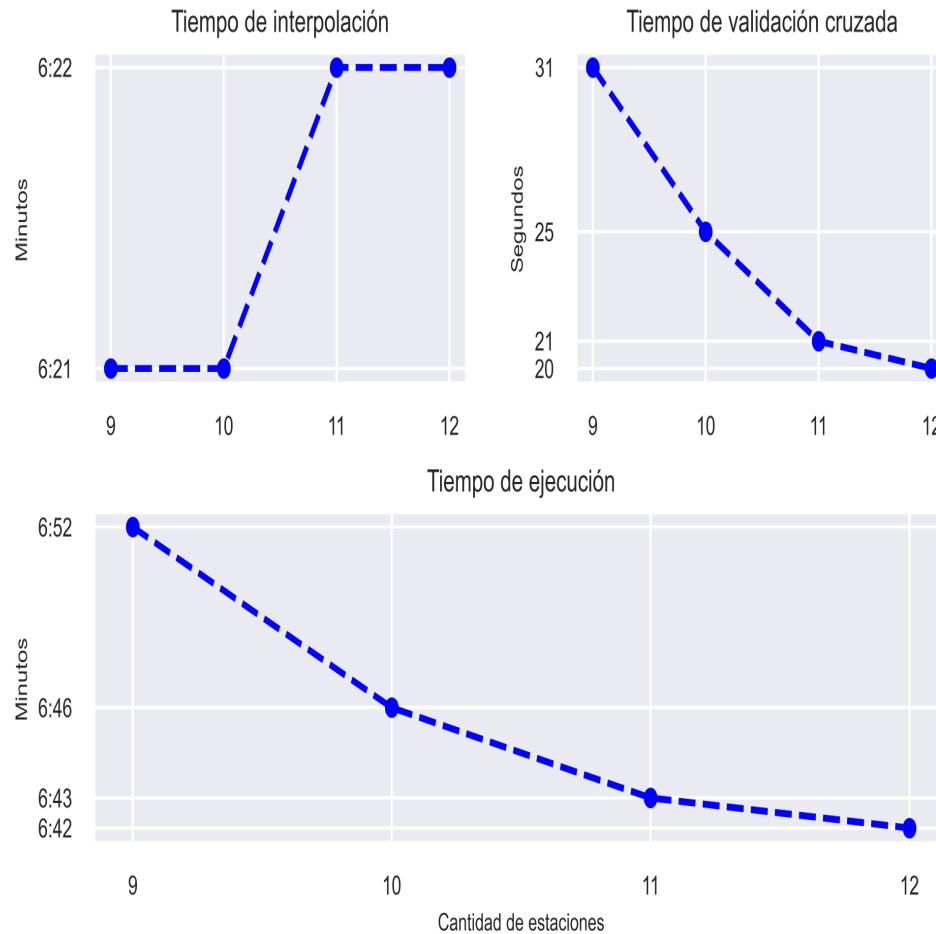


FIGURA 5.62: Tiempos de ejecución

En la figura 5.62 se muestran los tiempos de ejecución para los cuatro escenarios: nueve, diez, once y doce estaciones seleccionadas para interpolar: cuatro, tres, dos y una estación respectivamente. Se puede ver que los tiempos de ejecución decrecen conforme se aumenta el número de estaciones seleccionadas. Primero, se ve que el tiempo de ejecución de la interpolación aumenta muy poco conforme aumenta la cantidad de estaciones seleccionadas, de cinco horas con 21 minutos a cinco horas con 22 minutos. Después, se observa en los tiempos de ejecución de la validación cruzada, que éstos decrecen con mayor intensidad que lo que crecen los tiempos de ejecución de interpolación. En conclusión, podemos ver que el tiempo de ejecución decrece conforme aumenta la cantidad

de estaciones, esto debido a que los tiempos para ejecutar la validación cruzada influyen más que los tiempos de ejecución de las interpolaciones.

CAPÍTULO 6

RESULTADOS Y CONCLUSIONES

Del análisis antes visto en el capítulo anterior se eliminaron cinco del total de las quince variables, las cuales son: *presión barométrica*, *precipitación pluvial*, *humedad relativa*, *radiación solar* y *temperatura ambiental*. Esto debido a que en algunos registros de estas variables las mediciones registrarón el mismo valor en las trece estaciones, por lo que se vuelve imposible o poco adecuado hacer una interpolación, ya que los métodos geoespaciales necesitan que el valor mínimo reportado sea mayor estricto que el valor máximo reportado, por ejemplo:

1. **Deterministas:** en el caso de las Interpolación de Voronoi, que consiste en generar particiones del espacio de tal forma que cada partición corresponde a la estación seleccionada más cercana; cuando todas las estaciones seleccionadas presentan la misma cantidad de nivel de contaminante o variable, esto resultará en que toda la región tiene el mismo valor en cada punto interpolado. También para el caso de los métodos de interpolación de Función de Base Radial, los supuestos de calcular la interpolación a un punto desconocido, el cual depende de la distancia y del valor de las estaciones seleccionadas, no será una buena interpolación, pues el método necesita que los valores entre estaciones sean distintos para poder establecer como el valor y la distancia influyen en el nuevo

punto y se imposibilita si dos estaciones a diferente distancia reportan el mismo valor.

2. **Geoestadísticos:** Para el caso de los métodos de Kriging Ordinario y Kriging Universal, los supuestos piden que los valores sean distintos, pues los métodos infieren que existe cierta probabilidad que conlleve a que existe una distribución de probabilidad en los datos, si los valores son los mismos, se entiende que no existe incertidumbre en las mediciones, por lo que se pide que el valor superior de los valores reales observados sea menor estricto que el valor menor observado de los valores reales observados.

Se realizan 26,304 iteraciones correspondientes a las horas durante el período de años comprendido entre 2016 y 2018. Después, se realiza el experimento para cuatro instancias, de las cuales se escojen aleatoriamente un número de estaciones, luego se interpola sobre la región y se compara el valor obtenido de la interpolación contra el valor real en la posición geográfica de las estaciones interpoladas. Las instancias son del tipo:

- 9 estaciones seleccionadas y 4 estaciones interpoladas
- 10 estaciones seleccionadas y 2 estaciones interpoladas
- 11 estaciones seleccionadas y 3 estaciones interpoladas
- 12 estaciones seleccionadas y 1 estación interpolada

La experimentación se hizo considerando los cuatro escenarios antes mencionados y se evalúan once métodos de interpolación (TV, DIP, FBR M, FBR I, FBR G, FBR L, FBR C, FBR Q, FBR TPS, KO y KU) para las diez variables seleccionadas (CO, NO, NO₂, NO_x, O₃, PM₁₀, PM_{2.5}, SO₂, *velocidad del viento* y *dirección del viento*).

6.1. Variables del *Índice de AIRE y SALUD*

6.1.1 MONÓXIDO DE CARBONO (CO)

TABLA 6.1: CO: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.05	0.94	0.70	1.25	0.94	1.57
DIP	0.88	0.88	0.66	1.12	0.84	1.26
FBR M	1.00×10^{17}	1.33×10^{17}	1.00×10^{17}	1.11×10^{18}	8.32×10^{17}	1.23×10^{35}
FBR IM	1.14×10^{14}	2.62×10^{14}	1.97×10^{14}	4.92×10^{16}	3.69×10^{16}	2.42×10^{33}
FBR G	1.26×10^{16}	1.22×10^{16}	9.20×10^{15}	3.36×10^{17}	2.52×10^{17}	1.13×10^{35}
FBR L	1.09	1.00	0.75	1.36	1.02	1.85
FBR C	6.32×10^{17}	6.11×10^{17}	4.59×10^{17}	2.37×10^{18}	1.78×10^{18}	5.64×10^{36}
FBR Q	2.54×10^{18}	1.81×10^{18}	1.36×10^{18}	4.09×10^{18}	3.07×10^{18}	1.67×10^{37}
FBR TPS	6.78×10^{15}	4.20×10^{15}	3.15×10^{15}	1.97×10^{17}	1.47×10^{17}	3.88×10^{34}
KO	0.87	0.88	0.66	1.11	0.83	1.25
KU	1.13	1.00	0.75	1.45	1.08	2.10

De la tabla 6.1, en la cual se utilizan nueve estaciones para interpolar cuatro más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos, pero KO es mejor que DIP, ya que en los errores MAPE, RMSE, RMSEP y MSE obtiene mejores resultados, mientras que para el resto son iguales a los errores de DIP. En la figura 6.1, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

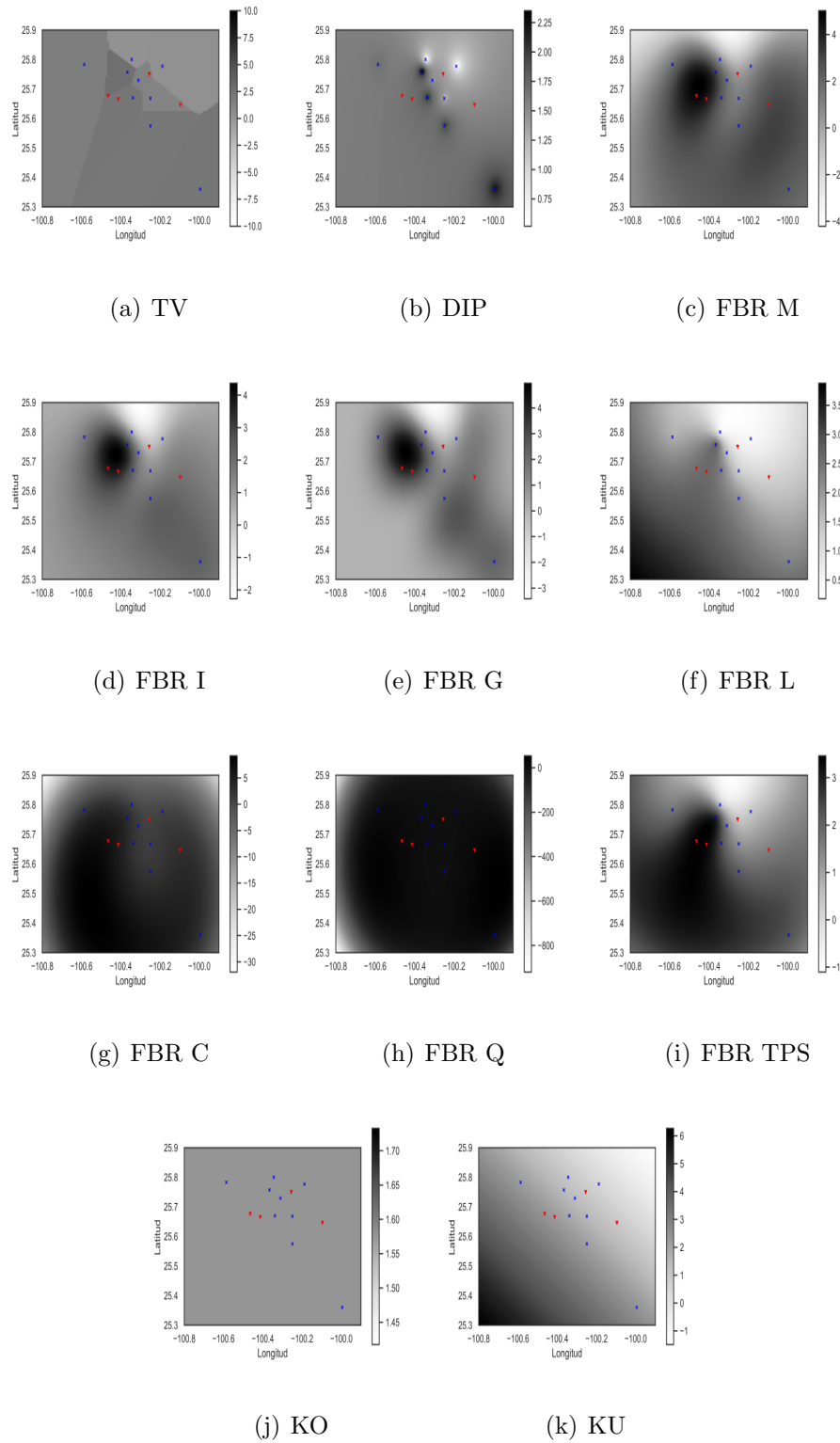


FIGURA 6.1: Interpolaciones de CO para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.2: CO: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.064	0.94	0.71	1.24	0.93	1.55
DIP	0.88	0.88	0.66	1.11	0.84	1.25
FBR M	1.35×10^{17}	1.88×10^{17}	1.41×10^{17}	1.31×10^{18}	9.92×10^{17}	1.73×10^{36}
FBR IM	6.25×10^{13}	1.16×10^{14}	8.79×10^{13}	3.28×10^{16}	2.47×10^{16}	1.07×10^{33}
FBR G	9.04×10^{15}	8.88×10^{15}	6.68×10^{15}	2.86×10^{17}	2.15×10^{17}	8.19×10^{34}
FBR L	1.09	1.00	0.75	1.32	0.99	1.75
FBR C	6.38×10^{17}	6.06×10^{17}	4.56×10^{17}	2.36×10^{18}	1.77×10^{18}	5.59×10^{36}
FBR Q	2.80×10^{18}	1.94×10^{18}	1.46×10^{18}	4.23×10^{18}	3.18×10^{18}	1.79×10^{37}
FBR TPS	1.90×10^{15}	2.10×10^{15}	1.58×10^{14}	1.39×10^{17}	1.04×10^{17}	1.94×10^{34}
KO	0.87	0.88	0.66	1.11	0.83	1.23
KU	1.14	1.00	0.75	1.41	1.06	1.99

De la tabla 6.2, en la cual se utilizan diez estaciones para interpolar otras tres, podemos ver que los métodos que obtienen peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción del método lineal (FBR L). Entre los métodos deterministas, el DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; en general DIP y KO son mejores que el resto de los métodos pero KO es mejor que DIP, ya que obtuvo resultados menores en los errores MAPE, RMSEP y MSE, mientras que en el resto fueron iguales a los errores de DIP. En la figura 6.2, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

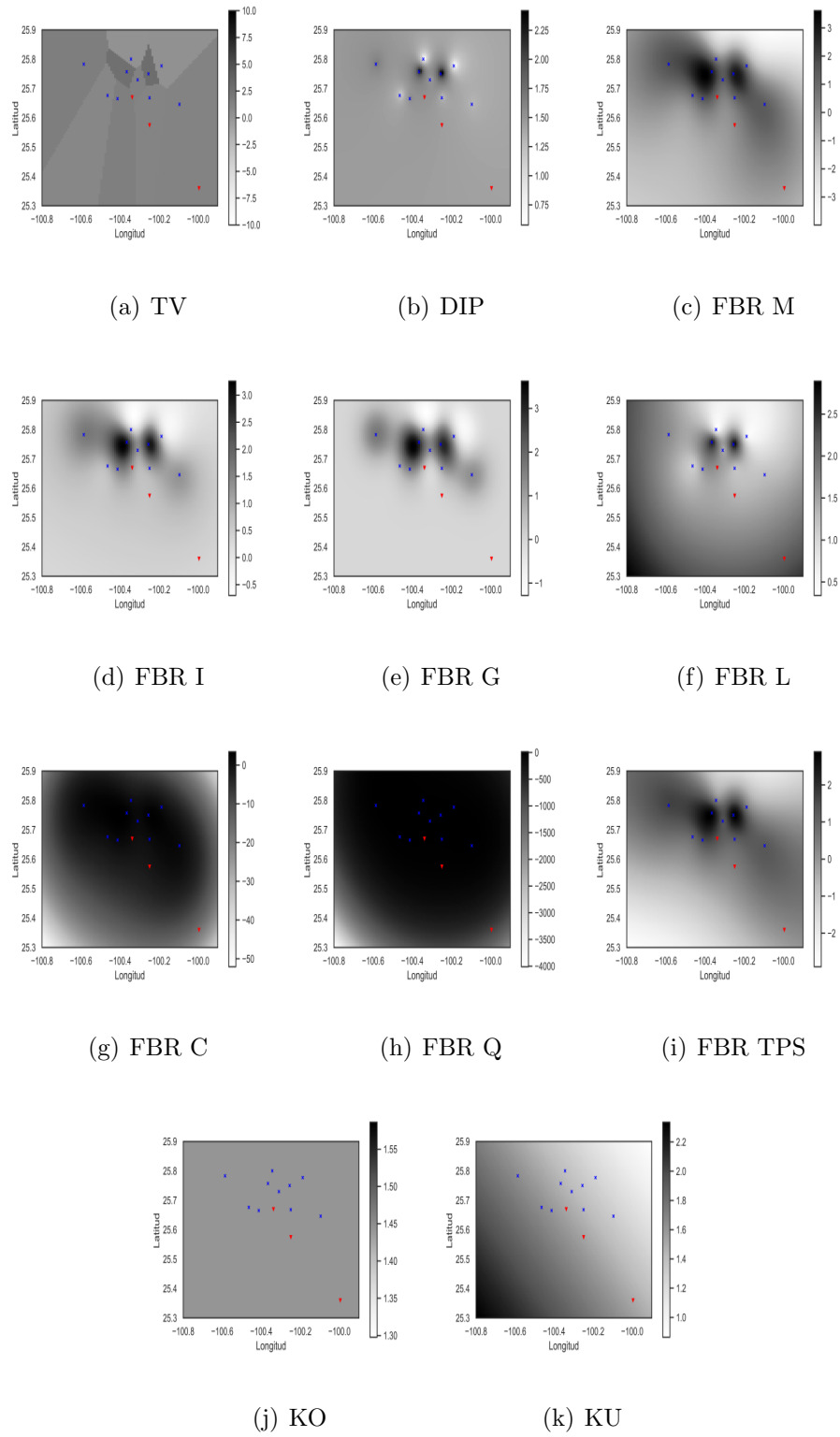


FIGURA 6.2: Interpolaciones de CO para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.3: CO: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.07	0.94	0.70	1.25	0.93	1.56
DIP	0.89	0.89	0.67	1.12	0.84	1.25
FBR M	1.79×10^{17}	2.51×10^{17}	1.88×10^{17}	1.52×10^{18}	1.14×10^{18}	2.31×10^{36}
FBR IM	1.23	1.16	0.87	1.65	1.24	2.72
FBR G	6.06×10^{16}	5.25×10^{15}	3.95×10^{15}	2.20×10^{17}	1.65×10^{17}	4.85×10^{34}
FBR L	1.10	1.01	0.76	1.32	0.99	1.75
FBR C	5.92×10^{17}	5.94×10^{17}	4.46×10^{17}	2.34×10^{18}	1.75×10^{18}	5.48×10^{36}
FBR Q	2.04×10^{18}	1.52×10^{18}	1.14×10^{18}	3.75×10^{18}	2.82×10^{18}	1.41×10^{37}
FBR TPS	3.71×10^{14}	3.50×10^{15}	2.63×10^{14}	5.68×10^{16}	4.27×10^{16}	3.23×10^{33}
KO	0.87	0.88	0.66	1.11	0.83	1.23
KU	1.11	0.98	0.74	1.34	1.01	1.82

De la tabla 6.3, en la cual se utilizan once estaciones para interpolar dos más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático y lineal (FBR I y FBR L). Entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; en general DIP y KO son mejores que el resto de los métodos pero KO es mejor que DIP ya que obtuvo en los errores MAPE, MAE, MAEP, RMSE, RMSEP y MSE mejores resultados que los de DIP. En la figura 6.3, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

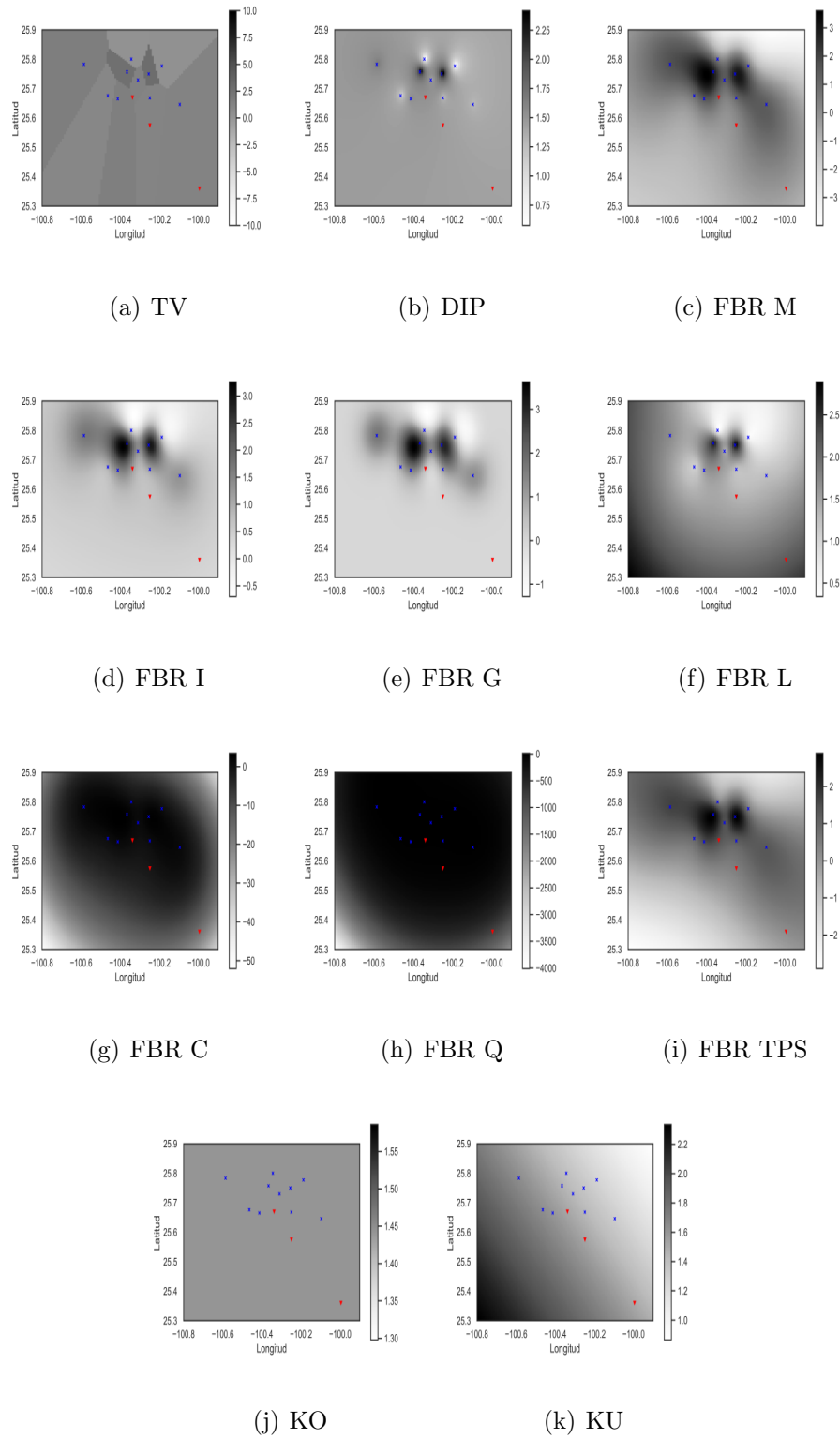


FIGURA 6.3: Interpolaciones de CO para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.4: CO: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.04	0.94	0.70	1.23	0.92	1.53
DIP	0.87	0.90	0.67	1.12	0.84	1.27
FBR M	2.20×10^{17}	3.29×10^{17}	2.46×10^{17}	1.74×10^{18}	1.30×10^{18}	3.04×10^{36}
FBR IM	1.12	1.14	0.85	1.60	1.19	2.56
FBR G	4.39×10^{15}	2.10×10^{15}	1.57×10^{14}	1.39×10^{17}	1.04×10^{17}	1.94×10^{34}
FBR L	1.09	1.03	0.77	1.34	1.00	1.80
FBR C	5.26×10^{17}	5.99×10^{17}	4.48×10^{17}	2.35×10^{18}	1.17×10^{18}	5.53×10^{36}
FBR Q	1.58×10^{18}	1.33×10^{18}	9.97×10^{17}	3.50×10^{18}	2.62×10^{18}	1.23×10^{37}
FBR TPS	1.31	1.22	0.91	1.70	1.27	2.90
KO	0.84	0.88	0.66	1.11	0.83	1.24
KU	1.07	0.98	0.73	1.32	0.99	1.76

De la tabla 6.4, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático, lineal y *thin plate splines* (FBR I, FBR L y FBR TPS), entre los métodos deterministas, el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero KO es mejor que DIP ya que en todos los errores son menores o igual a los errores de DIP. En la figura 6.4, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general, mientras se aumenta el número de estaciones para interpolar la variable CO, éstas bajan de forma rápida sus errores en algunos métodos como las Funciones de Base Radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios suceden con menor intensidad.

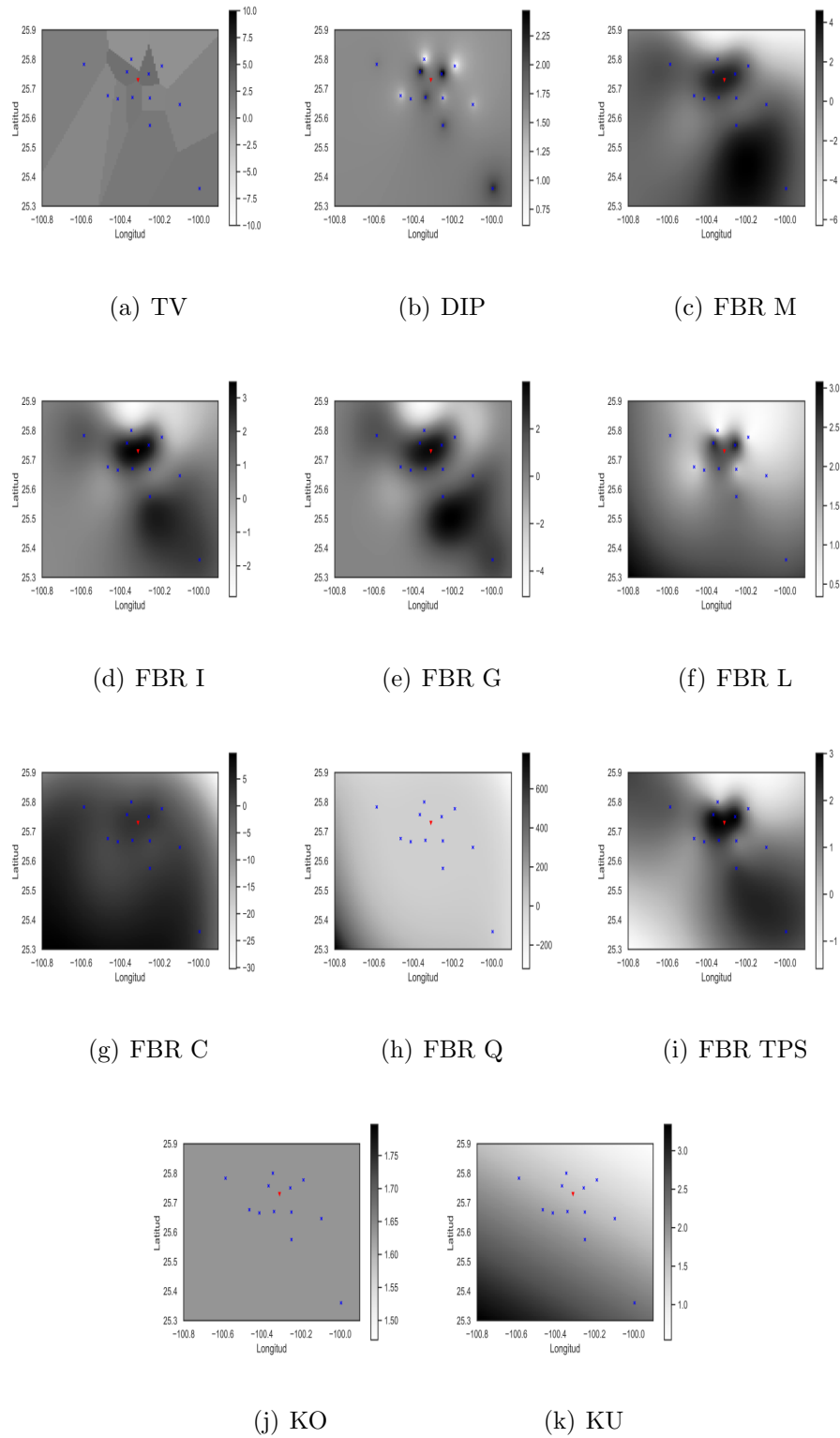


FIGURA 6.4: Interpolaciones de CO para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.1.2 DIÓXIDO DE NITRÓGENO (NO₂)TABLA 6.5: NO₂: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.79	6.02	0.57	9.42	0.90	88.84
DIP	0.60	4.49	0.43	7.59	0.72	57.64
FBR M	2.46×10^{14}	2.27×10^{14}	218×10^{14}	1.44×10^{17}	1.39×10^{16}	2.10×10^{34}
FBR IM	0.81	6.73	0.64	10.03	0.96	100.72
FBR G	6.26×10^{13}	2.62×10^{14}	2.52×10^{13}	4.92×10^{16}	4.72×10^{15}	2.42×10^{33}
FBR L	2.13	13.29	1.27	54.76	5.25	2,999.60
FBR C	1.20×10^{17}	6.57×10^{17}	6.31×10^{16}	2.46×10^{18}	2.36×10^{17}	6.06×10^{36}
FBR Q	2.78×10^{17}	1.64×10^{18}	1.57×10^{17}	3.88×10^{18}	3.72×10^{17}	1.50×10^{37}
FBR TPS	1.95	12.46	1.19	98.09	9.41	9,621.95
KO	0.63	4.66	0.44	7.87	0.75	61.98
KU	0.83	6.26	0.60	14.13	1.35	199.79

De la tabla 6.5, en la cual se utilizan nueve estaciones para interpolar cuatro más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático, lineal y *thin plate splines* (FBR I, FBR L y FBR TPS), entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en los errores MAPE, MAE, MAEP, RMSE, RMSEP y MSE encontró mejores resultados que los errores de KO. En la figura 6.5, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

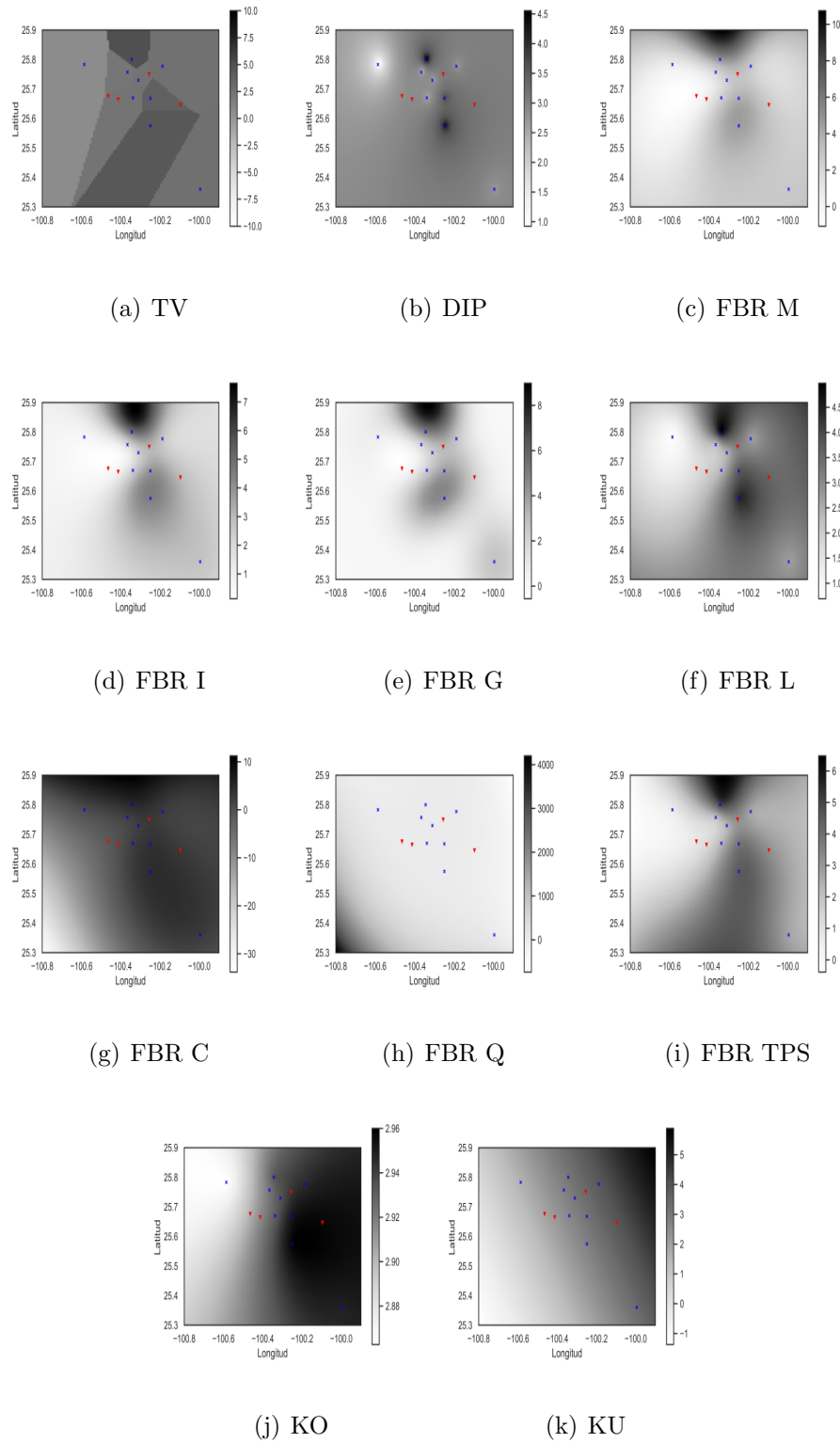


FIGURA 6.5: Interpolaciones de NO_2 para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.6: NO₂: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.81	6.05	0.58	9.45	0.90	89.38
DIP	0.60	4.46	0.42	7.57	0.72	57.32
FBR M	3.01×10^{14}	2.22×10^{15}	2.13×10^{14}	1.43×10^{17}	1.37×10^{16}	2.04×10^{34}
FBR IM	0.85	6.88	0.66	10.14	0.97	102.84
FBR G	1.32×10^{14}	4.67×10^{14}	4.49×10^{13}	6.56×10^{16}	6.31×10^{15}	4.31×10^{33}
FBR L	1.87	11.55	1.11	35.55	3.42	1,264.48
FBR C	1.00×10^{17}	5.51×10^{17}	5.30×10^{16}	2.25×10^{18}	2.16×10^{17}	5.08×10^{36}
FBR Q	2.50×10^{17}	1.55×10^{18}	1.49×10^{17}	3.77×10^{18}	3.63×10^{17}	1.42×10^{37}
FBR TPS	1.70	10.73	1.03	50.15	4.82	2,515.03
KO	0.63	4.62	0.44	7.84	0.75	61.55
KU	0.81	6.05	0.58	13.44	1.29	180.77

De la tabla 6.6 en la cual se utilizan diez estaciones para interpolar otras tres, podemos ver que los métodos que obtiene peores resultados de predicción, son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático, lineal y *thin plate splines* (FBR I, FBR L y FBR TPS), entre los métodos deterministas, el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en los errores MAPE, MAE, MAEP, RMSE RMSEP y MSE obtuvo mejores resultados que el método KO. En la figura 6.6, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

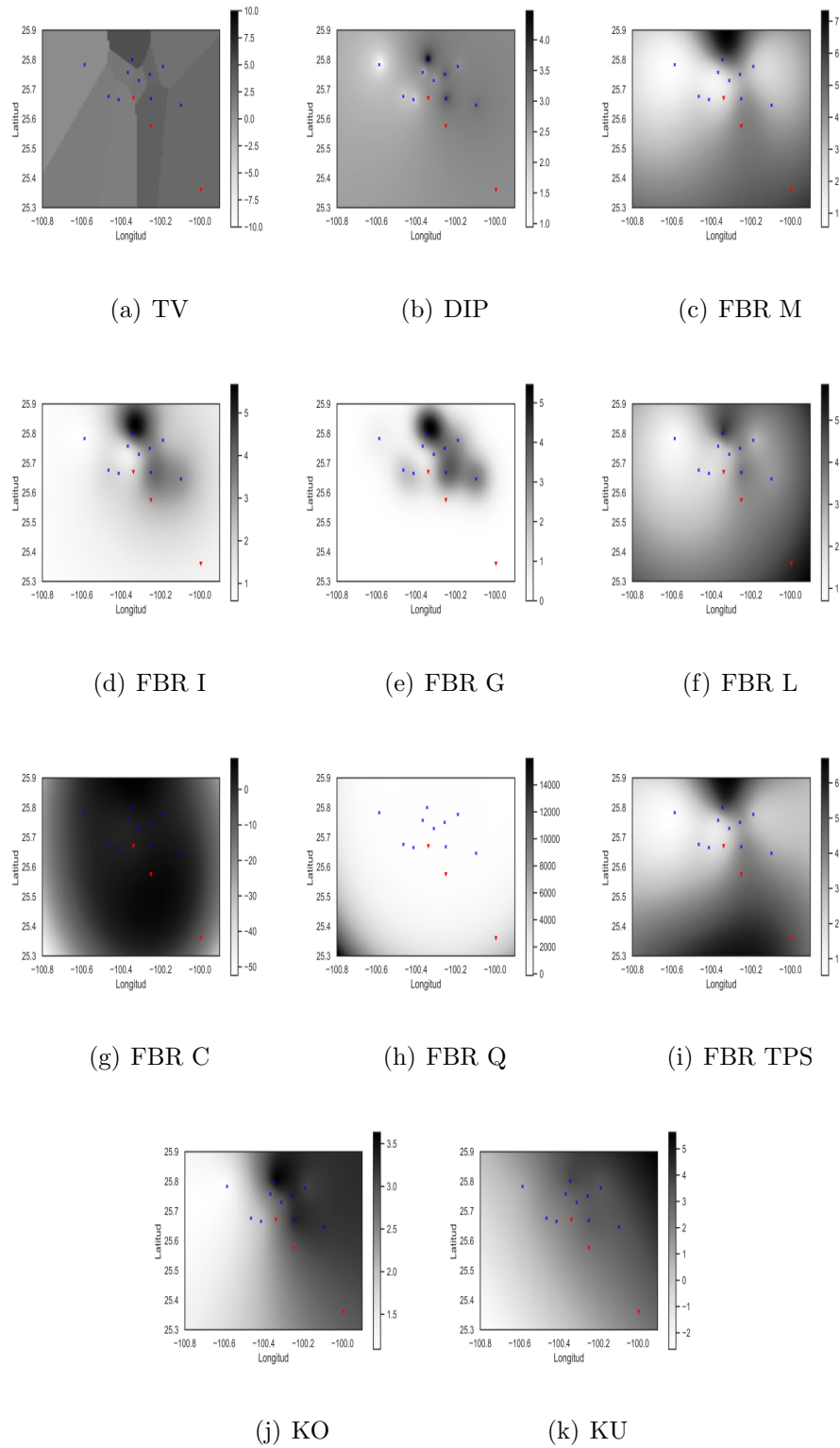


FIGURA 6.6: Interpolaciones de NO_2 para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.7: NO₂: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.82	6.10	0.58	9.49	0.91	90.14
DIP	0.60	4.44	0.42	7.54	0.72	56.97
FBR M	2.66×10^{14}	1.92×10^{15}	1.85×10^{14}	1.33×10^{17}	1.28×10^{16}	1.7×10^{34}
FBR IM	0.90	7.09	0.68	10.34	0.99	107.03
FBR G	5.59×10^{13}	3.50×10^{14}	3.37×10^{13}	5.68×10^{16}	5.47×10^{15}	3.23×10^{33}
FBR L	1.70	10.38	0.99	26.98	2.59	728.01
FBR C	8.81×10^{16}	4.77×10^{17}	4.59×10^{16}	2.09×10^{18}	2.01×10^{17}	4.40×10^{36}
FBR Q	1.70×10^{17}	1.05×10^{18}	1.01×10^{17}	3.11×10^{18}	2.99×10^{17}	9.68×10^{36}
FBR TPS	1.52	9.56	0.92	35.03	3.37	1,227.34
KO	0.629	4.59	0.44	7.77	0.74	60.50
KU	0.78	5.78	0.55	11.58	1.11	134.26

De la tabla 6.7, en la cual se utilizan once estaciones para interpolar dos estaciones más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático, lineal y *thin plate splines* (FBR I, FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geostatísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos, pero DIP es mejor que KO ya que en todos los errores son menores a los errores de KO. En la figura 6.7, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

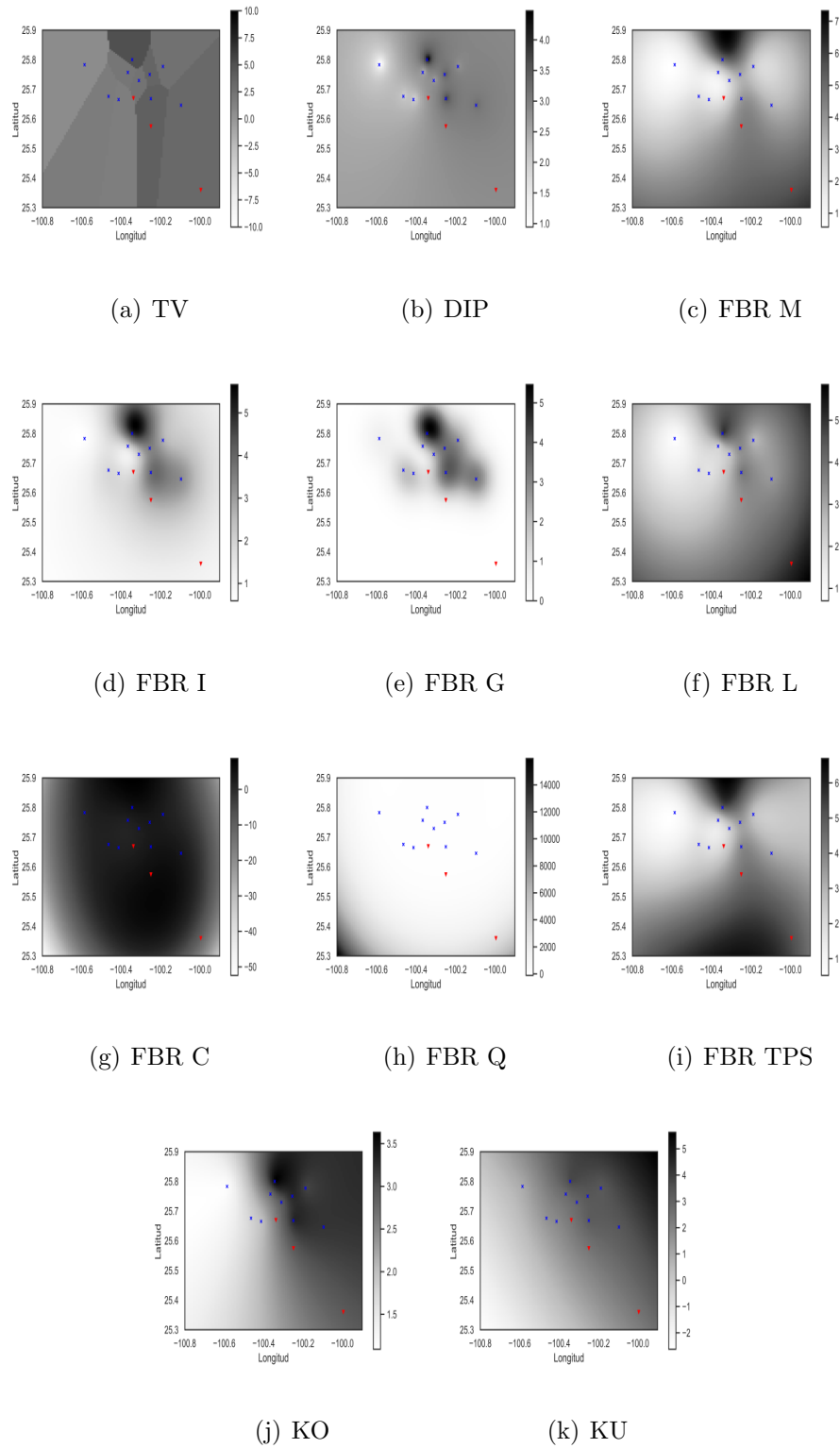


FIGURA 6.7: Interpolaciones de NO_2 para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.8: NO₂: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.82	6.11	0.58	9.44	0.90	89.19
DIP	0.59	4.42	0.42	7.53	0.72	56.79
FBR M	2.52×10^{14}	2.80×10^{15}	2.68×10^{14}	1.60×10^{17}	1.53×10^{16}	2.58×10^{34}
FBR IM	0.93	7.30	0.69	10.43	0.99	108.96
FBR G	1.20	10.27	0.98	19.31	1.84	373.16
FBR L	1.54	9.48	0.90	19.73	1.88	389.62
FBR C	7.40×10^{16}	4.32×10^{17}	4.14×10^{16}	1.99×10^{18}	1.91×10^{17}	3.99×10^{36}
FBR Q	1.20×10^{17}	8.15×10^{17}	7.80×10^{16}	2.74×10^{18}	2.62×10^{17}	7.51×10^{36}
FBR TPS	1.30	8.38	0.80	16.85	1.61	284.10
KO	0.61	4.57	0.43	7.76	0.74	60.24
KU	0.74	5.65	0.54	10.42	0.99	108.76

De la tabla 6.8, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático, lineal y *thin plate splines* (FBR I, FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos, pero DIP es mejor que KO ya que todos los errores encontrados por DIP son menores a los errores encontrados por KO. En la figura 6.8, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general, mientras se aumenta el número de estaciones para interpolar la variable NO₂, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

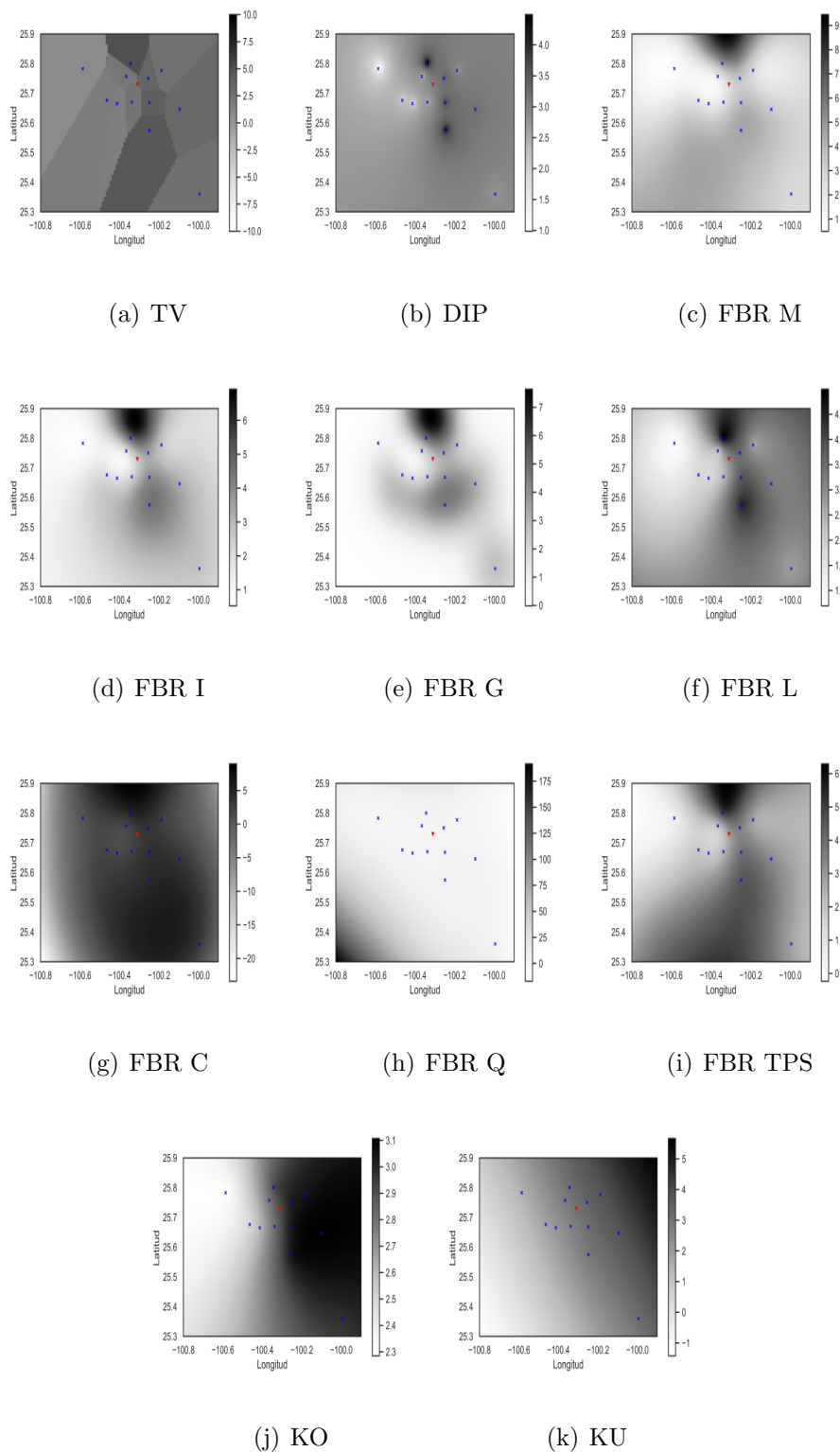


FIGURA 6.8: Interpolaciones de NO_2 para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.1.3 OZONO (O_3)TABLA 6.9: O_3 : 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.69	11.28	0.44	17.26	0.67	298.22
DIP	0.49	7.95	0.31	13.01	0.50	169.37
FBR M	5.53×10^{13}	5.39×10^{16}	2.11×10^{15}	7.05×10^{17}	2.76×10^{16}	4.97×10^{35}
FBR IM	5.66×10^{14}	5.61×10^{15}	2.19×10^{14}	2.27×10^{17}	8.90×10^{15}	5.17×10^{34}
FBR G	7.32×10^{15}	6.20×10^{16}	2.42×10^{15}	7.56×10^{17}	2.96×10^{16}	5.72×10^{35}
FBR L	1.21	23.71	0.92	58.65	2.29	3,440.80
FBR C	5.96×10^{16}	9.94×10^{17}	3.89×10^{16}	3.02×10^{18}	1.18×10^{17}	9.17×10^{36}
FBR Q	1.50×10^{17}	1.95×10^{18}	7.67×10^{16}	4.25×10^{18}	1.66×10^{17}	1.80×10^{37}
FBR TPS	43.01×10^{12}	5.25×10^{14}	2.05×10^{13}	6.96×10^{16}	2.72×10^{13}	4.85×10^{33}
KO	0.51	8.17	0.32	3.37	0.52	178.88
KU	1.18×10^{13}	1.30×10^{16}	5.11×10^{14}	3.47×10^{17}	1.35×10^{16}	1.20×10^{35}

De la tabla 6.9, en la cual se utilizan nueve estaciones para interpolar cuatro más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los método lineal (FBR L), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos, el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores a excepción de RMSE obtuvo mejores resultados que los errores de KO. En la figura 6.9, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

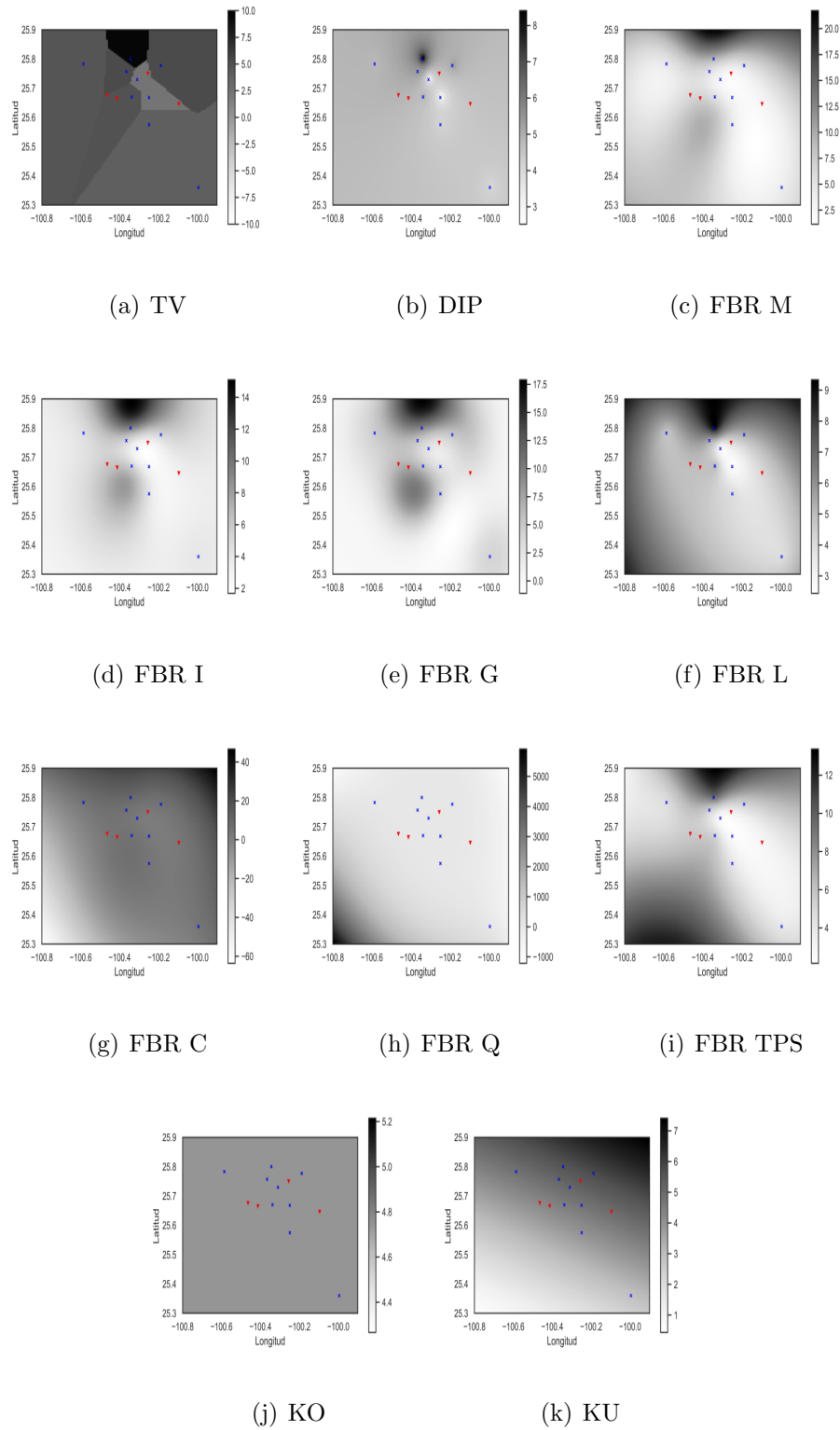


FIGURA 6.9: Interpolaciones de O_3 para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.10: O₃: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.69	11.43	0.44	17.52	0.68	307.02
DIP	0.48	7.91	0.30	13.01	0.50	169.39
FBR M	5.85×10^{15}	5.76×10^{16}	2.25×10^{15}	7.29×10^{17}	2.85×10^{16}	5.31×10^{35}
FBR IM	3.87×10^{14}	4.44×10^{15}	1.73×10^{14}	2.02×10^{17}	791×10^{13}	4.09×10^{34}
FBR G	8.71×10^{15}	7.67×10^{16}	3.00×10^{15}	8.41×10^{17}	3.29×10^{16}	7.08×10^{35}
FBR L	1.09	20.95	0.81	46.95	1.83	2,204.98
FBR C	5.33×10^{16}	8.96×10^{17}	3.50×10^{16}	2.87×10^{18}	1.12×10^{17}	8.26×10^{36}
FBR Q	1.41×10^{17}	1.71×10^{18}	6.69×10^{16}	3.97×10^{18}	1.55×10^{17}	1.57×10^{37}
FBR TPS	3.01×10^{13}	5.84×10^{13}	2.28×10^{13}	7.34×10^{16}	2.87×10^{15}	5.39×10^{33}
KO	0.50	8.12	0.31	13.36	0.52	178.62
KU	1.03×10^{15}	1.22×10^{16}	4.79×10^{14}	3.36×10^{17}	1.31×10^{16}	1.13×10^{35}

De la tabla 6.10, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los método lineal (FBR L), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, pues en todos los errores encontrados por DIP son menores a los errores de KO. En la figura 6.10 se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

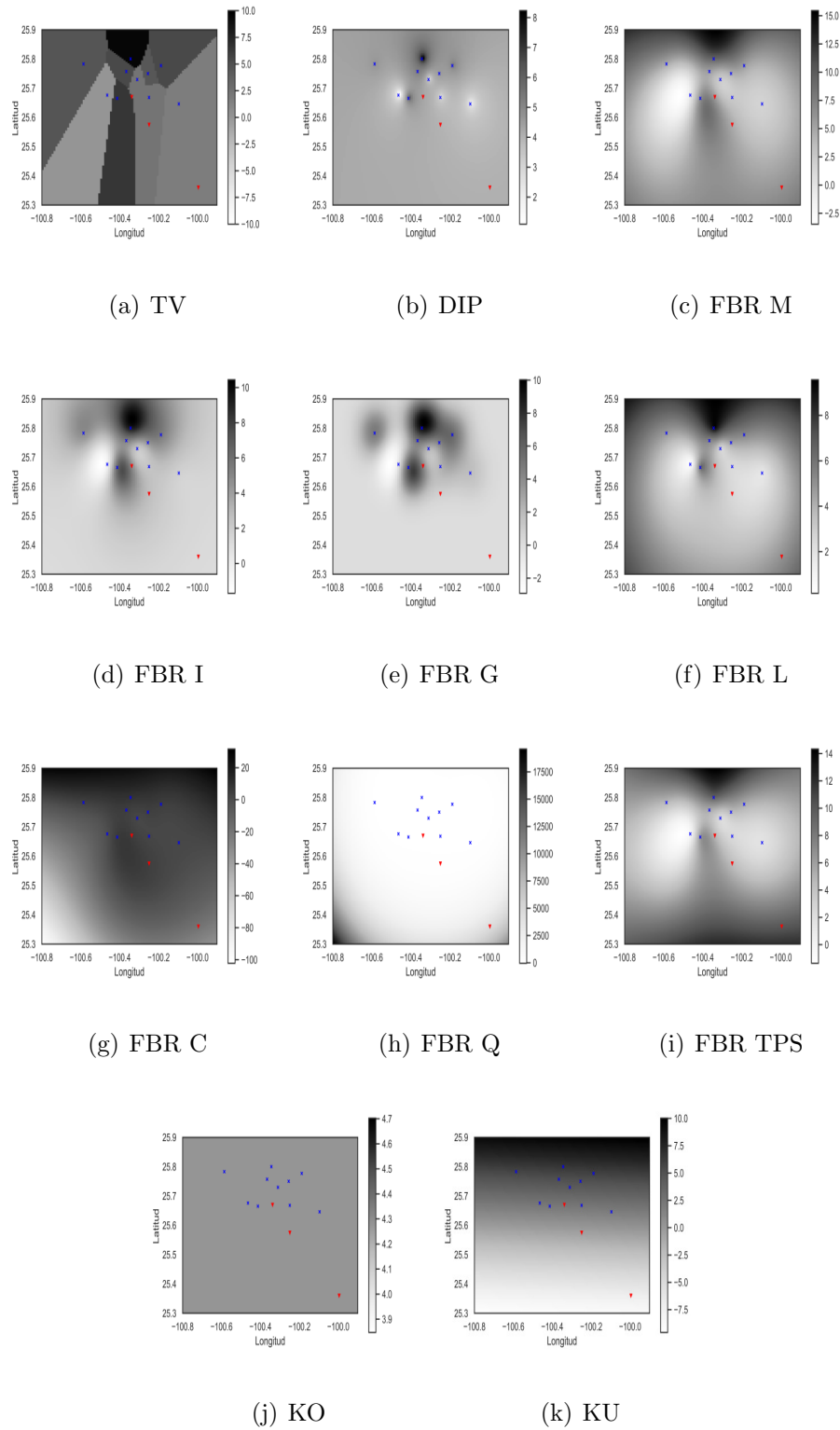


FIGURA 6.10: Interpolaciones de O_3 para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.11: O₃: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.70	11.48	0.44	17.64	0.69	311.32
DIP	0.48	7.86	0.30	12.94	0.50	167.62
FBR M	4.30×10^{15}	5.39×10^{16}	2.11×10^{15}	7.05×10^{17}	2.76×10^{16}	4.98×10^{35}
FBR IM	3.30×10^{15}	3.85×10^{15}	1.50×10^{14}	1.88×10^{17}	7.37×10^{15}	3.55×10^{34}
FBR G	9.63×10^{15}	9.11×10^{16}	3.56×10^{15}	9.16×10^{17}	3.58×10^{16}	8.40×10^{35}
FBR L	1.034	19.85	0.77	42.77	1.67	1,829.37
FBR C	4.65×10^{16}	8.23×10^{17}	3.22×10^{16}	2.75×10^{18}	1.07×10^{17}	7.59×10^{36}
FBR Q	1.00×10^{17}	1.18×10^{18}	4.64×10^{16}	3.31×10^{18}	1.29×10^{17}	1.09×10^{37}
FBR TPS	2.55×10^{13}	3.50×10^{14}	1.37×10^{13}	5.68×10^{16}	2.22×10^{15}	3.23×10^{33}
KO	0.50	8.061	0.315	13.27	0.519	176.16
KU	8.86×10^{14}	1.06×10^{16}	4.18×10^{14}	3.14×10^{17}	1.22×10^{16}	9.86×10^{34}

De la tabla 6.11, en la cual se utilizan once estaciones para interpolar dos más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los método lineal (FBR L), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores son menores a los errores de KO. En la figura 6.11, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

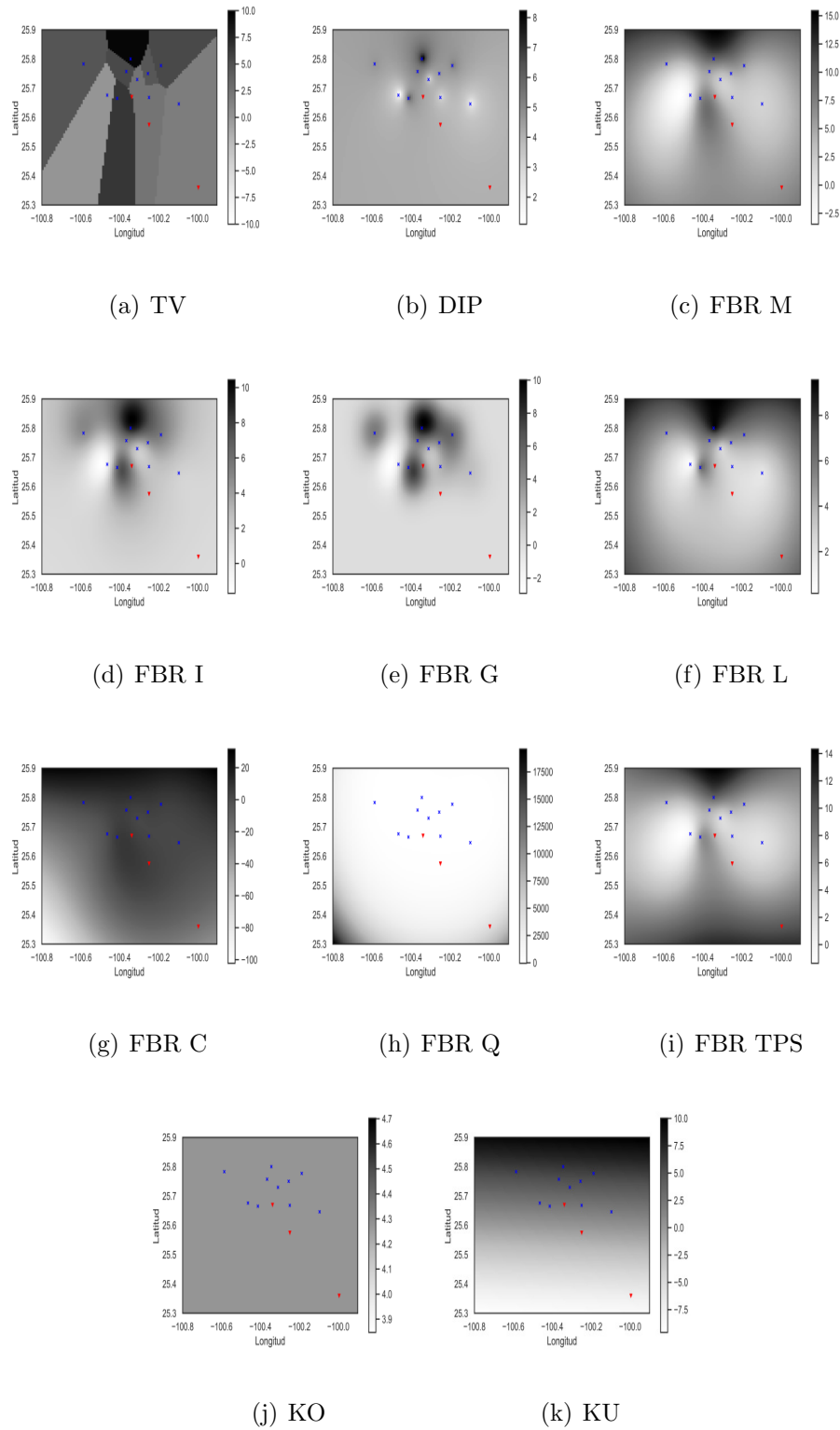


FIGURA 6.11: Interpolaciones de O_3 para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.12: O₃: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.72	11.71	0.46	18.03	0.70	325.16
DIP	0.48	7.80	0.30	12.92	0.50	167.16
FBR M	4.13×10^{15}	5.54×10^{16}	2.17×10^{15}	7.14×10^{17}	2.81×10^{16}	5.10×10^{35}
FBR IM	7.22×10^{14}	7.36×10^{15}	2.89×10^{14}	2.60×10^{17}	1.02×10^{16}	6.79×10^{34}
FBR G	1.31×10^{16}	1.38×10^{17}	5.44×10^{15}	1.13×10^{18}	4.44×10^{16}	1.27×10^{36}
FBR L	0.99	18.76	0.73	38.54	1.51	1,485.79
FBR C	4.38×10^{16}	7.94×10^{17}	3.12×10^{16}	2.70×10^{18}	1.06×10^{17}	7.32×10^{36}
FBR Q	6.95×10^{16}	8.07×10^{17}	3.17×10^{16}	2.72×10^{18}	1.07×10^{17}	7.44×10^{36}
FBR TPS	0.91	15.56	0.61	29.04	1.14	843.58
KO	0.50	7.97	0.31	13.18	0.51	173.73
KU	9.14×10^{14}	1.05×10^{16}	4.13×10^{14}	3.11×10^{17}	1.22×10^{16}	9.70×10^{34}

De la tabla 6.12, en la cual se utilizan doce estaciones para interpolar una estación más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos lineal y thin plate splines (FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores obtenidos por DIP son menores que los errores de KO. En la figura 6.12, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general, mientras se aumenta el número de estaciones para interpolar la variable O₃, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

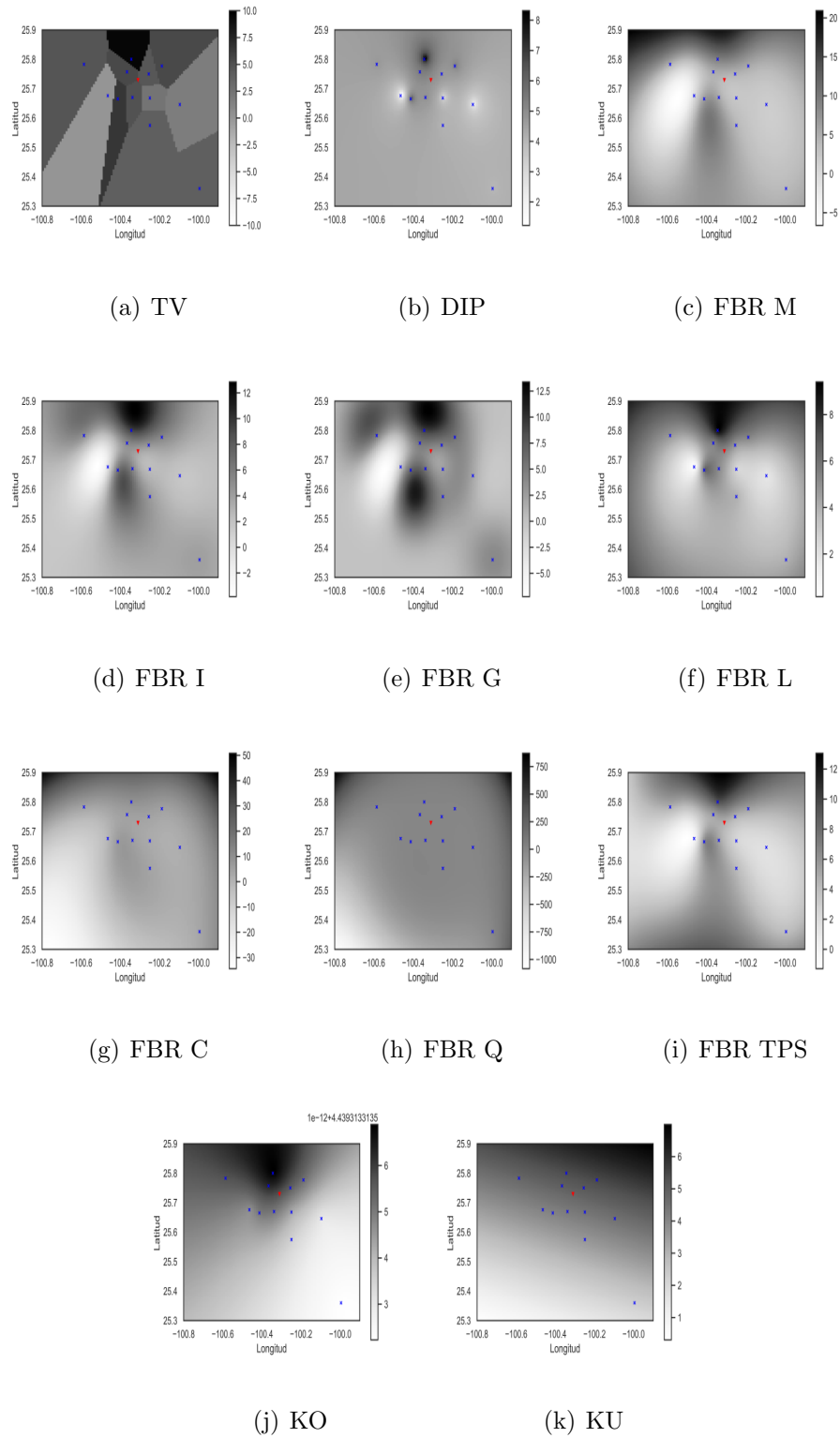


FIGURA 6.12: Interpolaciones de O_3 para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.1.4 PARTÍCULAS MENORES A 10 MICRAS (PM_{10})TABLA 6.13: PM_{10} : 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.53	26.72	0.44	44.30	0.73	1,963.15
DIP	0.39	19.11	0.31	32.78	0.54	1,074.85
FBR M	1.70×10^{12}	3.94×10^{13}	$6,50 \times 10^{11}$	1.90×10^{17}	3.14×10^{13}	3.63×10^{34}
FBR IM	0.66	33.85	0.55	58.90	0.97	3,470.28
FBR G	1.83×10^{12}	4.82×10^{13}	7.95×10^{11}	2.10×10^{17}	3.47×10^{13}	4.44×10^{34}
FBR L	2.28	118.00	1.94	527.01	8.69	277,744.08
FBR C	2.68×10^{16}	9.91×10^{17}	1.63×10^{16}	3.02×10^{18}	4.98×10^{16}	9.14×10^{36}
FBR Q	4.76×10^{16}	1.72×10^{18}	2.85×10^{16}	3.98×10^{18}	6.57×10^{16}	1.58×10^{37}
FBR TPS	1.30	63.84	1.05	366.53	6.04	134,350.51
KO	0.38	19.44	0.32	33.35	0.55	1,112.30
KU	0.48	24.23	0.39	42.66	0.70	1,820.70

De la tabla 6.13, en la cual se utilizan nueve estaciones para interpolar otras cuatro, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores son menores a los errores de KO a excepción del RMSE donde KO es menor a DIP. En la figura 6.13, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

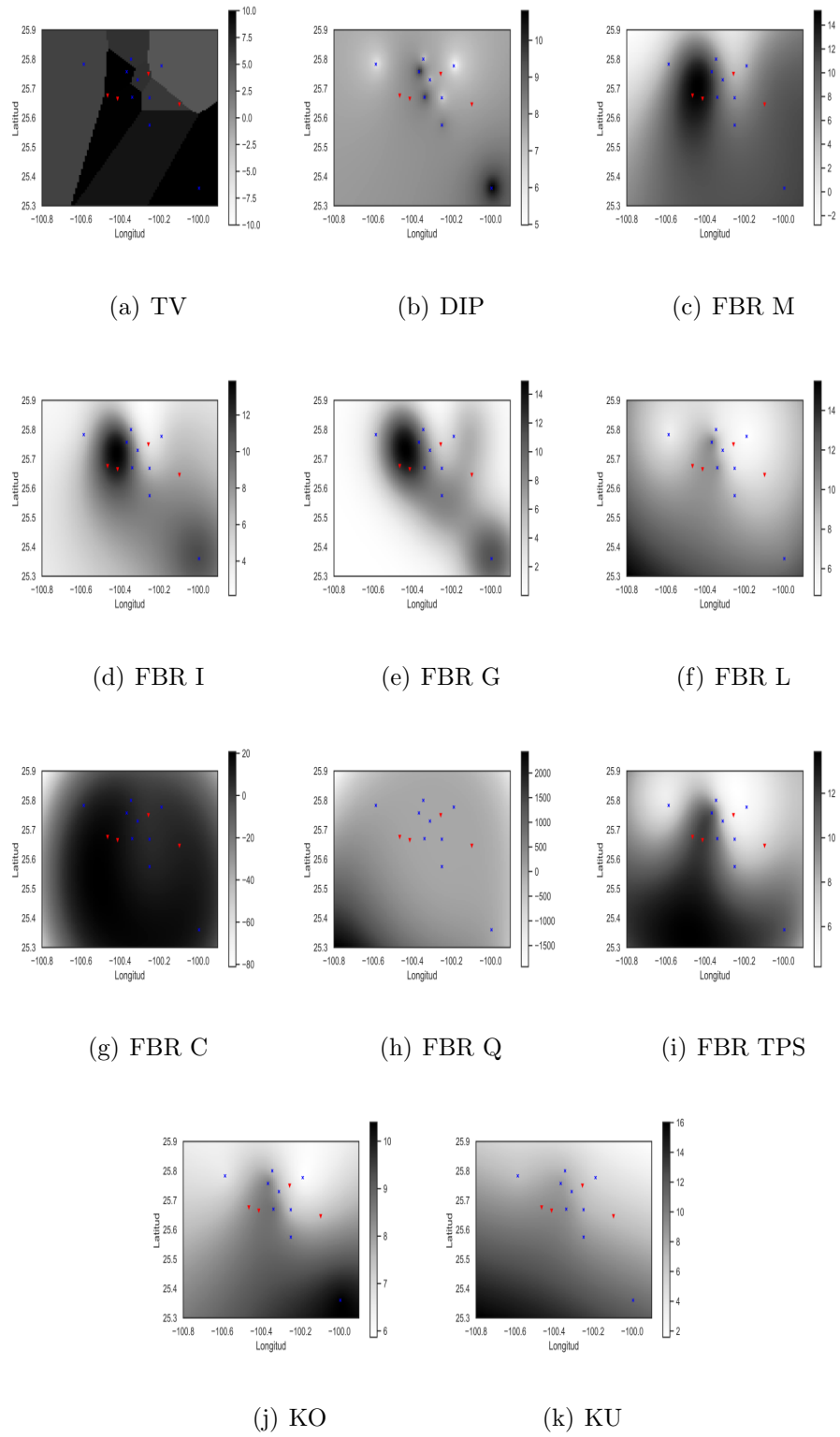


FIGURA 6.13: Interpolaciones de PM_{10} para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.14: PM₁₀: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.52	26.64	0.43	44.14	0.72	1,949.08
DIP	0.39	18.92	0.31	32.66	0.53	1,067.22
FBR M	2.71×10^{14}	5.96×10^{15}	9.82×10^{13}	2.34×10^{17}	3.86×10^{15}	5.49×10^{34}
FBR IM	0.68	34.40	0.56	59.96	0.98	3,595.99
FBR G	1.44×10^{14}	4.55×10^{15}	7.51×10^{13}	2.05×10^{17}	3.37×10^{15}	4.20×10^{34}
FBR L	1.8	96.12	1.58	389.29	6.41	151546.85
FBR C	2.27×10^{16}	8.73×10^{17}	1.43×10^{16}	2.83×10^{18}	4.67×10^{16}	8.05×10^{36}
FBR Q	4.15×10^{16}	1.47×10^{18}	42.42×10^{16}	3.67×10^{18}	6.06×10^{16}	1.35×10^{37}
FBR TPS	1.1	53.83	0.88	300.63	4.95	90,383.85
KO	0.38	19.20	0.31	33.06	0.54	1,092.97
KU	0.47	23.43	0.38	41.41	0.68	1,715.16

De la tabla 6.14, en la cual se utilizan diez estaciones para interpolar tres más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general, DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, pues en los errores MAE RMSEP y MSE obtuvo mejores resultados que KO. En la figura 6.14, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

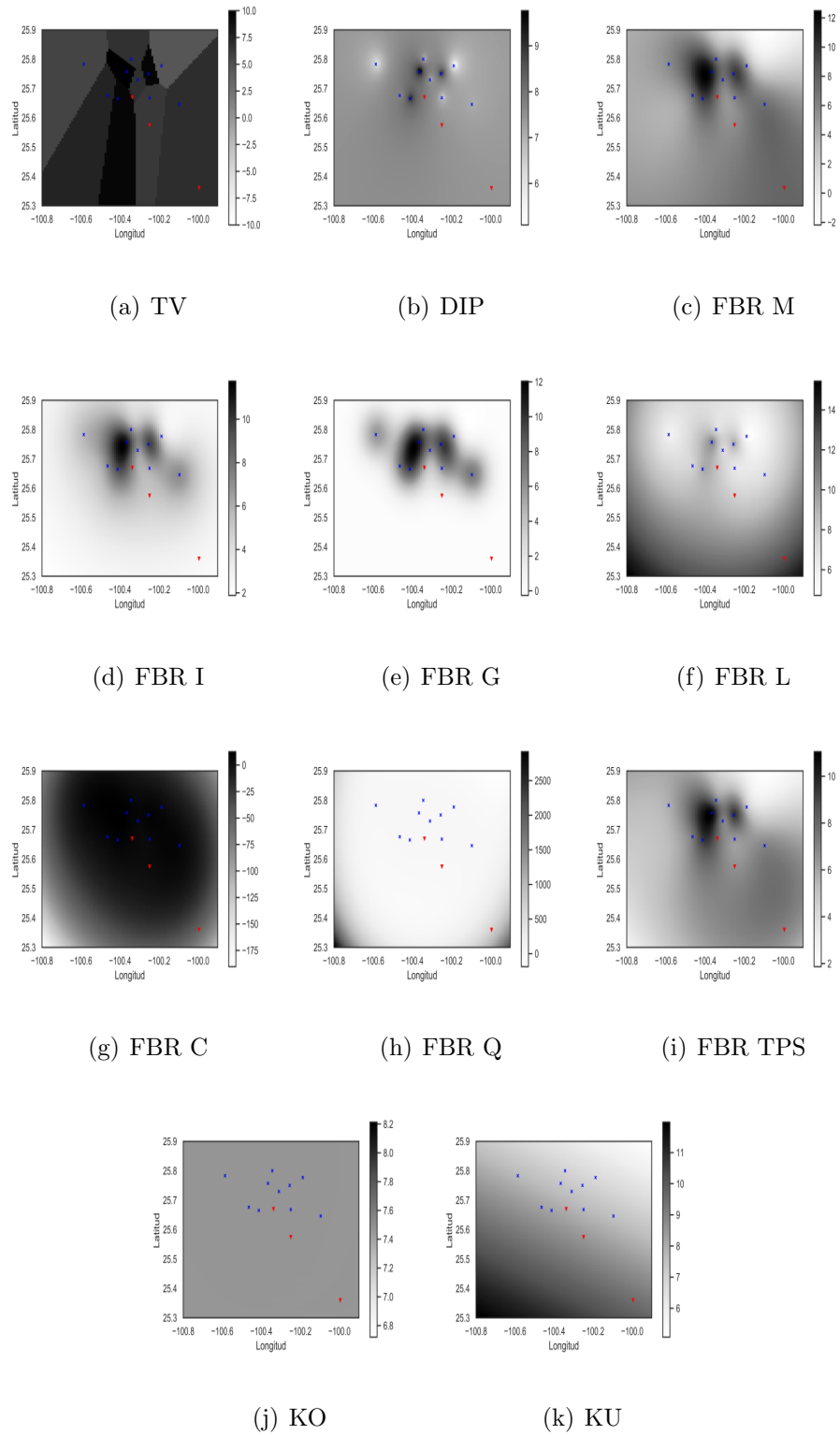


FIGURA 6.14: Interpolaciones de PM_{10} para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.15: PM₁₀: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.52	26.77	0.44	44.63	0.73	1,992.57
DIP	0.39	18.85	0.31	32.34	0.53	1,046.16
FBR M	4.50×10^{14}	8.06×10^{15}	1.32×10^{14}	2.72×10^{17}	4.48×10^{14}	7.43×10^{34}
FBR IM	0.70	35.27	0.58	62.89	1.03	3,955.48
FBR G	2.61×10^{14}	7.36×10^{15}	1.21×10^{14}	2.60×10^{17}	4.28×10^{15}	6.79×10^{34}
FBR L	1.61	81.48	1.34	249.97	4.11	62489.07
FBR C	2.14×10^{16}	8.05×10^{17}	1.32×10^{16}	2.72×10^{18}	4.48×10^{16}	7.43×10^{36}
FBR Q	2.95×10^{16}	1.00×10^{18}	1.65×10^{16}	3.04×10^{18}	5.01×10^{16}	9.28×10^{36}
FBR TPS	0.97	46.76	0.76	143.21	2.35	20,510.59
KO	0.38	19.10	0.31	32.76	0.53	1,073.31
KU	0.47	22.96	0.37	39.44	0.64	1,555.53

De la tabla 6.15, en la cual se utilizan once estaciones para interpolar dos más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores son menores a los errores de KO a excepción del RMSE donde KO es menor a DIP. En la figura 6.15, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

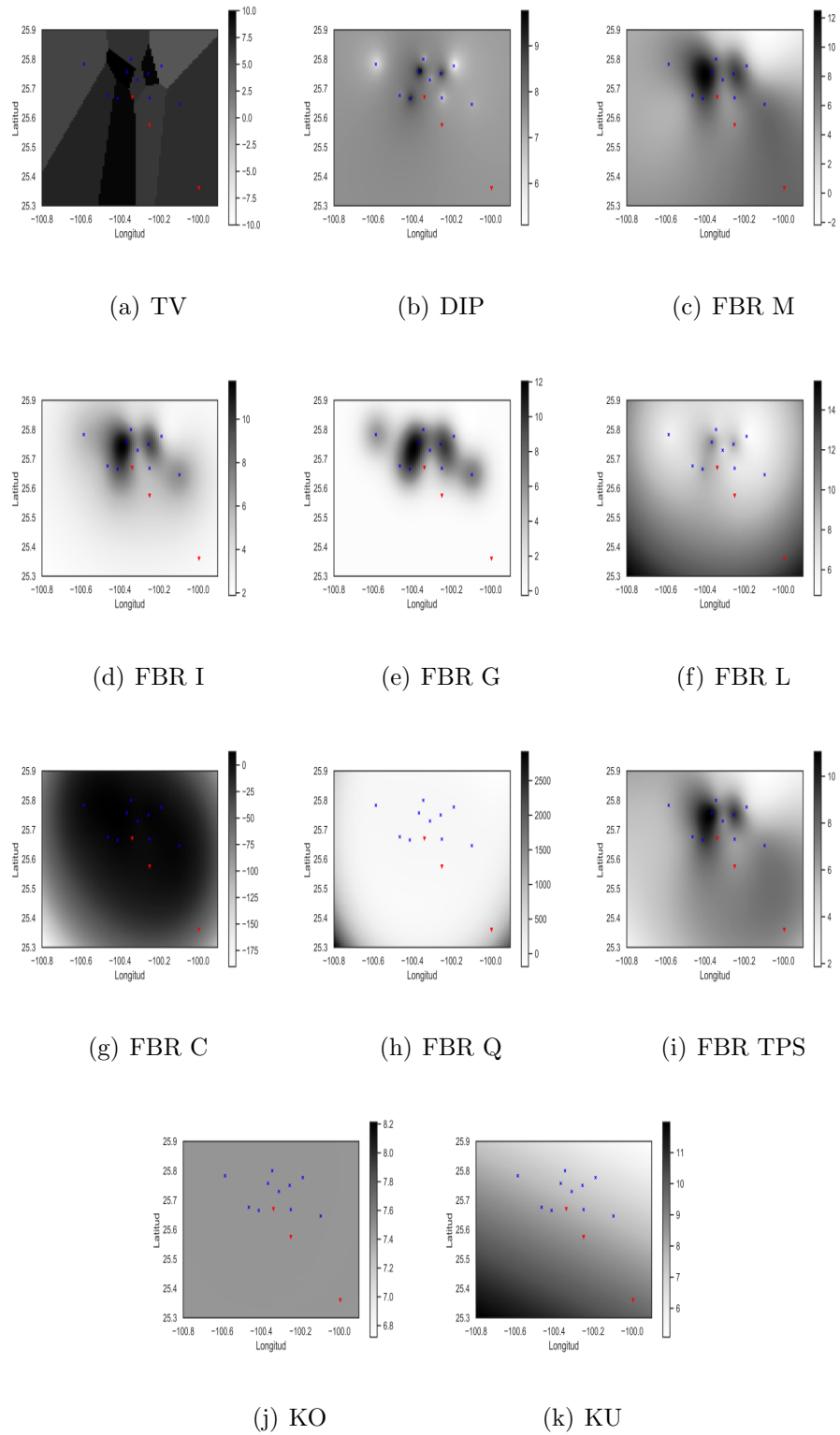


FIGURA 6.15: Interpolaciones de PM_{10} para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.16: PM₁₀: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.51	26.60	0.43	44.89	0.73	2,015.75
DIP	0.38	18.78	0.30	33.13	0.54	1,098.17
FBR M	6.47×10^{14}	9.11×10^{15}	1.49×10^{14}	2.89×10^{17}	4.76×10^{15}	8.40×10^{34}
FBR IM	0.71	35.72	0.58	60.68	0.99	3,682.12
FBR G	5.15×10^{14}	1.22×10^{16}	2.01×10^{14}	3.36×10^{17}	5.53×10^{15}	1.13×10^{35}
FBR L	1.42	71.04	1.16	201.42	3.31	40570.18
FBR C	1.98×10^{16}	7.58×10^{17}	1.24×10^{16}	2.64×10^{18}	4.34×10^{16}	6.99×10^{36}
FBR Q	2.36×10^{16}	7.61×10^{17}	1.25×10^{16}	2.65×10^{18}	4.35×10^{16}	7.02×10^{36}
FBR TPS	0.84	40.77	0.67	88.49	1.45	7,831.72
KO	0.37	19.03	0.31	33.51	0.55	1,123.50
KU	0.45	22.44	0.36	38.84	0.63	1,508.66

De la tabla 6.16, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos, pero DIP es mejor que KO. En la figura 6.16, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general mientras se aumenta el número de estaciones para interpolar la variable PM₁₀, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

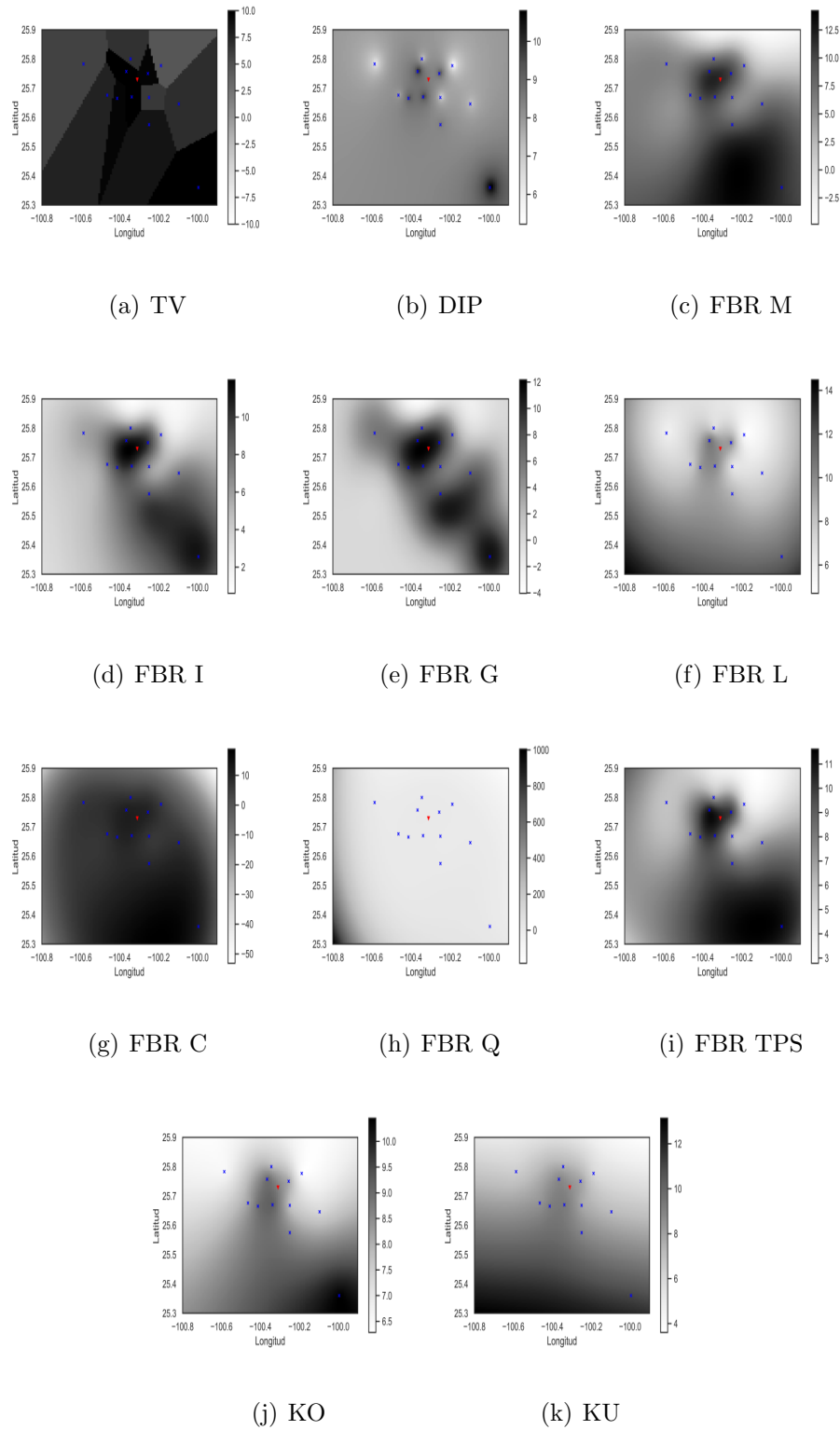


FIGURA 6.16: Interpolaciones de PM_{10} para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.1.5 PARTÍCULAS MENORES A 2.5 MICRAS (PM_{2.5})TABLA 6.17: PM_{2.5}: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.51	8.98	0.41	14.58	0.67	212.84
DIP	0.39	6.66	0.30	11.20	0.51	125.50
FBR M	3.01×10^{14}	4.38×10^{15}	2.01×10^{14}	2.01×10^{17}	9.25×10^{15}	4.04×10^{34}
FBR IM	0.62	11.80	0.54	18.91	0.87	357.84
FBR G	3.79×10^{14}	5.61×10^{15}	2.58×10^{14}	2.27×10^{17}	1.04×10^{16}	5.17×10^{34}
FBR L	1.47	24.36	1.12	71.36	3.28	5,092.49
FBR C	6.92×10^{16}	9.24×10^{17}	4.25×10^{16}	2.92×10^{18}	1.34×10^{17}	8.53×10^{36}
FBR Q	1.24×10^{17}	1.84×10^{18}	8.48×10^{16}	4.11×10^{18}	1.89×10^{17}	1.69×10^{37}
FBR TPS	0.76	1.00	2.13	16.69	46.33	2,146.72
KO	0.40	6.72	0.30	11.25	0.51	126.60
KU	0.50	8.53	0.39	14.27	0.65	203.82

De la tabla 6.17 en la cual se utilizan nueve estaciones para interpolar otras cuatro, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores son menores a los errores encontrados por KO. En la figura 6.17, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

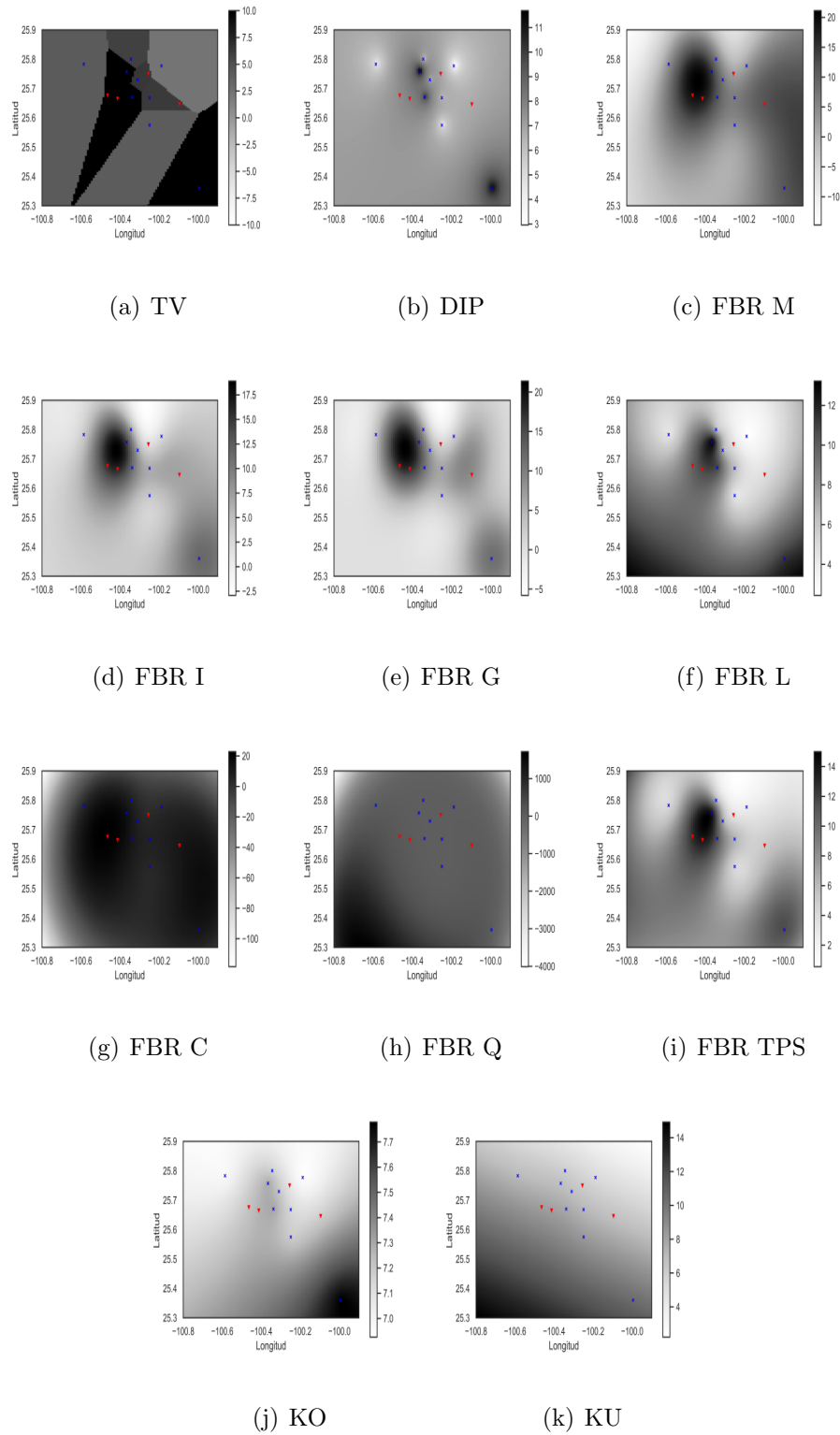


FIGURA 6.17: Interpolaciones de $PM_{2.5}$ para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.18: PM_{2.5}: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.50	8.87	0.40	14.73	0.67	217.14
DIP	0.39	6.65	0.30	11.91	0.54	141.91
FBR M	5.53×10^{14}	6.66×10^{15}	3.05×10^{14}	2.47×10^{17}	1.13×10^{16}	6.14×10^{34}
FBR IM	0.63	11.74	0.53	18.13	0.83	328.73
FBR G	6.50×10^{14}	1.13×10^{16}	5.20×10^{14}	3.23×10^{17}	1.48×10^{16}	1.04×10^{35}
FBR L	1.23	20.37	0.93	53.65	2.46	2,878.46
FBR C	5.92×10^{16}	7.75×10^{17}	3.56×10^{16}	2.67×10^{18}	1.22×10^{17}	7.15×10^{36}
FBR Q	1.10×10^{17}	1.62×10^{18}	7.47×10^{16}	3.87×10^{18}	1.77×10^{17}	1.49×10^{37}
FBR TPS	0.86	14.38	0.66	33.17	1.52	1,100.33
KO	0.39	6.73	0.30	11.98	0.55	143.55
KU	0.48	8.27	0.37	14.01	0.64	196.53

De la tabla 6.18, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO. En la figura 6.18, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

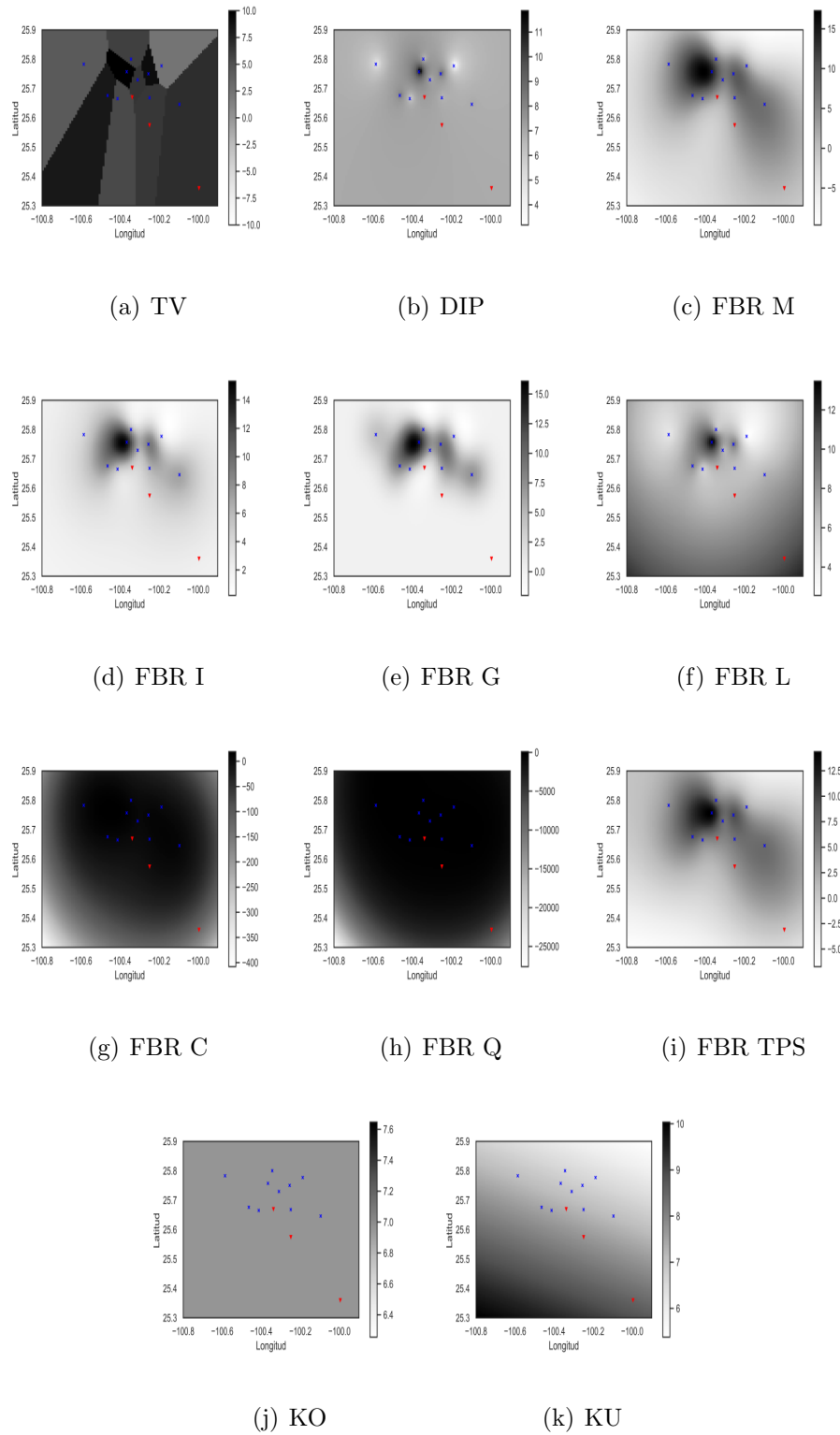


FIGURA 6.18: Interpolaciones de $PM_{2.5}$ para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.19: PM_{2.5}: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.49	8.63	0.39	13.75	0.63	189.25
DIP	0.39	6.55	0.30	10.97	0.50	120.38
FBR M	3.85×10^{14}	4.20×10^{15}	1.94×10^{14}	1.97×10^{17}	9.08×10^{15}	3.88×10^{34}
FBR IM	0.65	11.68	0.53	19.36	0.89	374.93
FBR G	5.00×10^{14}	7.36×10^{15}	3.39×10^{14}	2.60×10^{17}	1.20×10^{16}	6.79×10^{34}
FBR L	1.15	18.80	0.86	48.14	2.22	2,317.68
FBR C	5.17×10^{16}	6.52×10^{17}	3.00×10^{16}	2.45×10^{18}	1.13×10^{17}	6.01×10^{36}
FBR Q	8.39×10^{16}	1.13×10^{18}	5.25×10^{16}	3.24×10^{18}	1.49×10^{17}	1.05×10^{37}
FBR TPS	0.81	13.08	0.60	26.51	1.22	703.24
KO	0.40	6.65	0.30	11.03	0.50	121.88
KU	0.48	7.97	0.36	12.87	0.59	165.69

De la tabla 6.19, en la cual se utilizan once estaciones para interpolar dos estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE,. En general DIP y KO son mejores que el resto de los métodos, pero DIP es mejor que KO, pues en los errores MAPE, MAE, RMSE y MSE obtuvo mejores resultados que los obtenidos por KO. En la figura 6.19, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

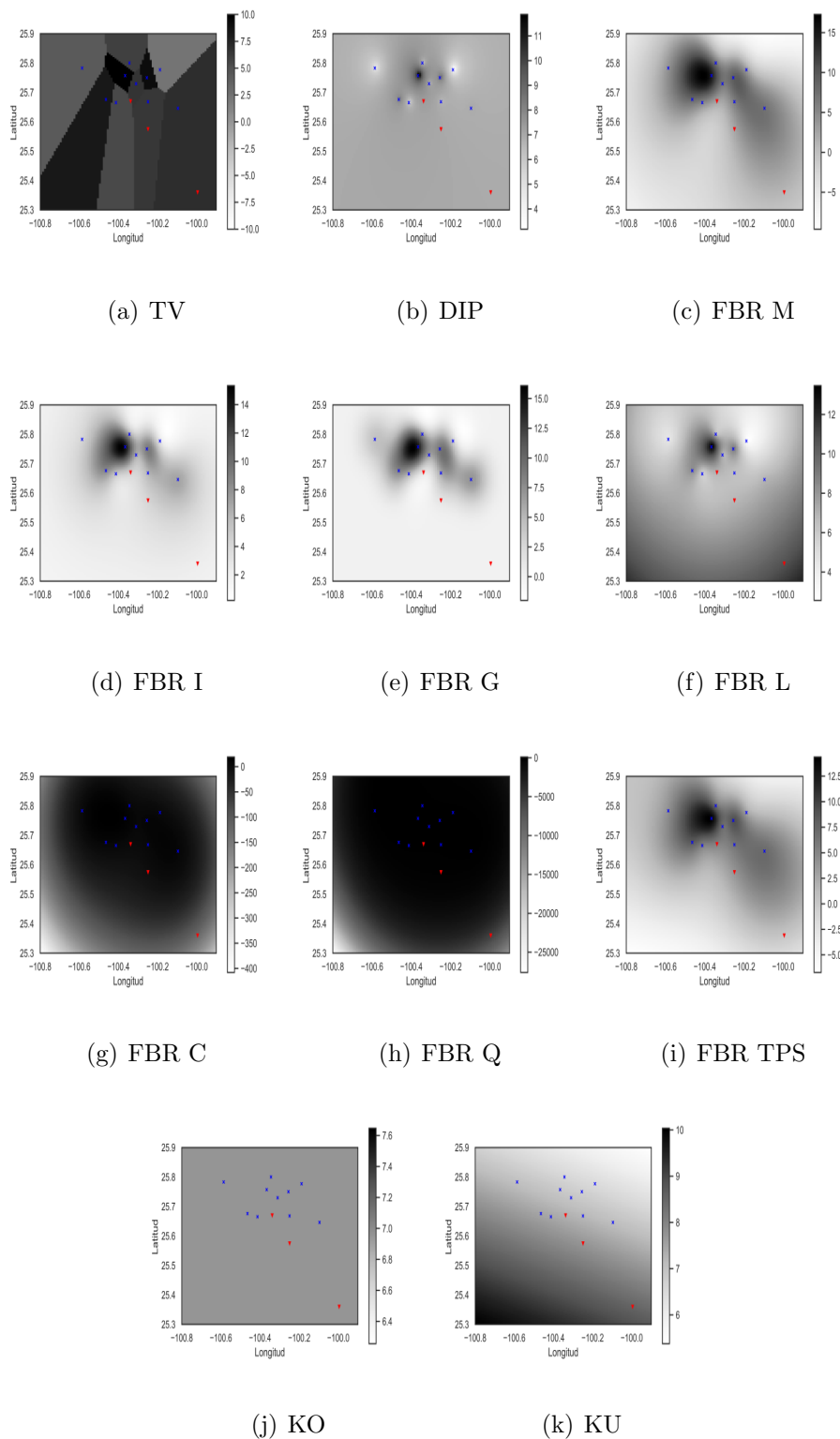


FIGURA 6.19: Interpolaciones de $PM_{2.5}$ para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.20: PM_{2.5}: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.48	8.50	0.39	14.51	0.66	210.61
DIP	0.38	6.50	0.29	11.90	0.54	141.82
FBR M	5.47×10^{12}	6.31×10^{15}	2.90×10^{14}	2.41×10^{17}	1.11×10^{16}	5.82×10^{34}
FBR IM	0.63	11.65	0.53	19.55	0.90	382.21
FBR G	1.09×10^{15}	1.22×10^{16}	5.65×10^{14}	3.36×10^{17}	1.55×10^{16}	1.13×10^{35}
FBR L	1.03	17.37	0.80	42.50	1.95	1,807.09
FBR C	4.55×10^{16}	5.92×10^{17}	2.72×10^{16}	2.33×10^{18}	1.07×10^{17}	5.46×10^{36}
FBR Q	6.80×10^{16}	8.67×10^{17}	4.00×10^{16}	2.82×10^{18}	1.30×10^{17}	8.00×10^{36}
FBR TPS	0.73	12.08	0.55	22.84	1.05	521.87
KO	0.39	6.61	0.30	11.98	0.55	143.61
KU	0.46	7.79	0.35	13.28	0.61	176.57

De la tabla 6.20, en la cual se utilizan doce estaciones para interpolar una estación más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos inverso multicuadrático lineal, *thin plate splines* (FBR IM, FBR L y FBR TPS); entre los métodos deterministas, DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geostatísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, ya que en todos los errores son menores a los errores encontrados por el método KO. En la figura 6.20, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general mientras se aumenta el número de estaciones para interpolar la variable PM_{2.5}, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

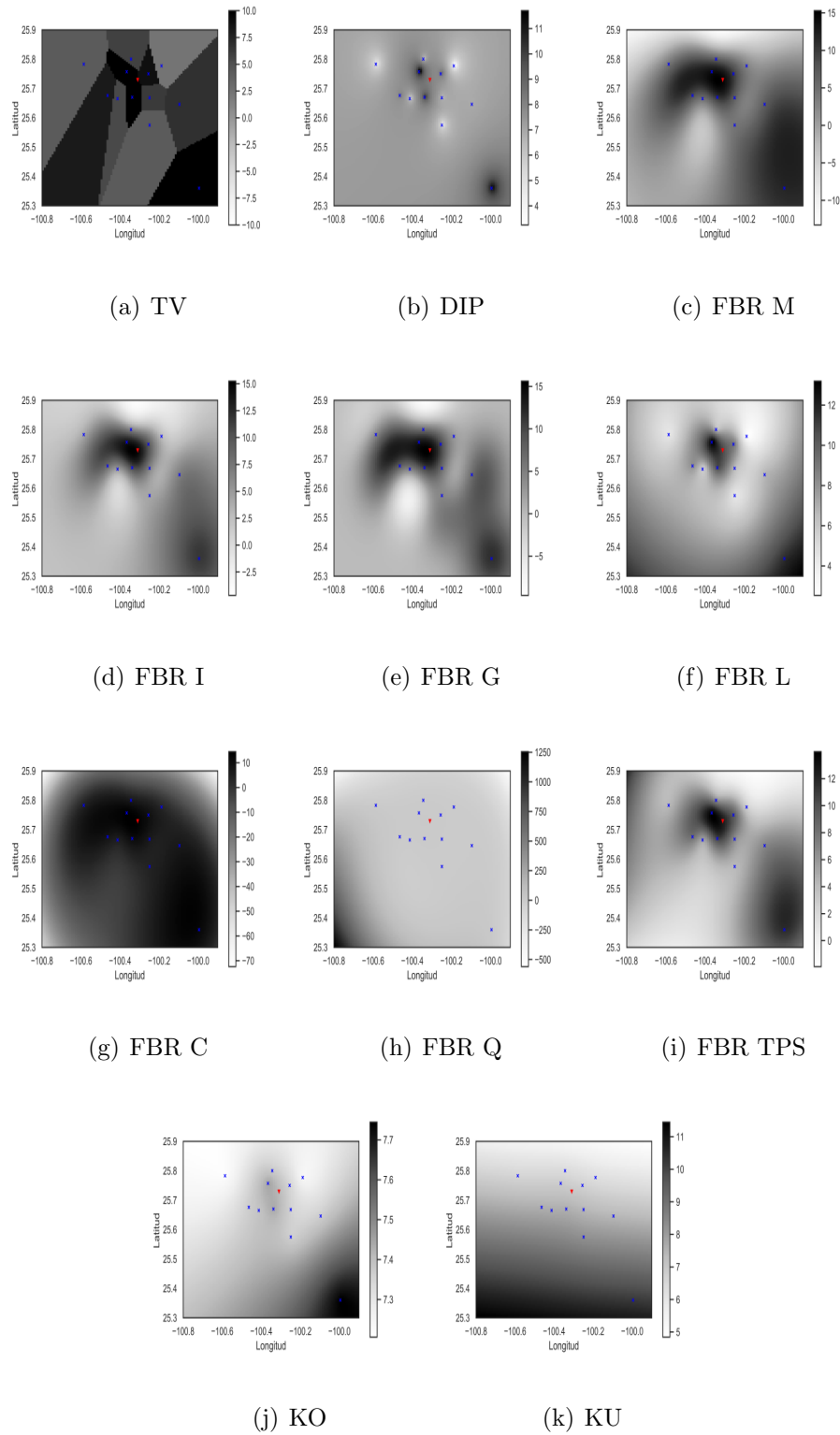


FIGURA 6.20: Interpolaciones de $PM_{2.5}$ para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.1.6 BIÓXIDO DE AZUFRE (SO₂)TABLA 6.21: SO₂: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.71	4.81	0.61	6.73	0.86	45.42
DIP	0.45	3.17	0.40	5.07	0.64	25.71
FBR M	4.80	31.64	4.05	1,156.98	148.15	1,338,608.74
FBR IM	0.88	6.03	0.77	11.80	1.51	139.43
FBR G	1.74	11.60	1.48	165.06	21.13	27,246.64
FBR L	2.36	13.37	1.71	65.28	8.35	4,261.71
FBR C	1.41×10^{14}	6.25×10^{14}	8.01×10^{13}	7.52×10^{16}	9.63×10^{14}	5.66×10^{33}
FBR Q	1.72×10^{17}	9.58×10^{17}	1.22×10^{17}	2.96×10^{18}	3.80×10^{17}	8.81×10^{36}
FBR TPS	1.42	8.69	1.11	43.64	5.58	1,904.94
KO	0.46	3.21	0.41	5.11	0.65	26.18
KU	0.59	3.91	0.50	7.54	0.96	56.92

De la tabla 6.21 en la cual se utilizan nueve estaciones para interpolar otras cuatro, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial (FBR C y FBR Q) de los cuales el mejor prediciendo es el método inverso multicuadrático (FBR IM); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, ya que en todos los errores son menores a los errores de KO. En la figura 6.21, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

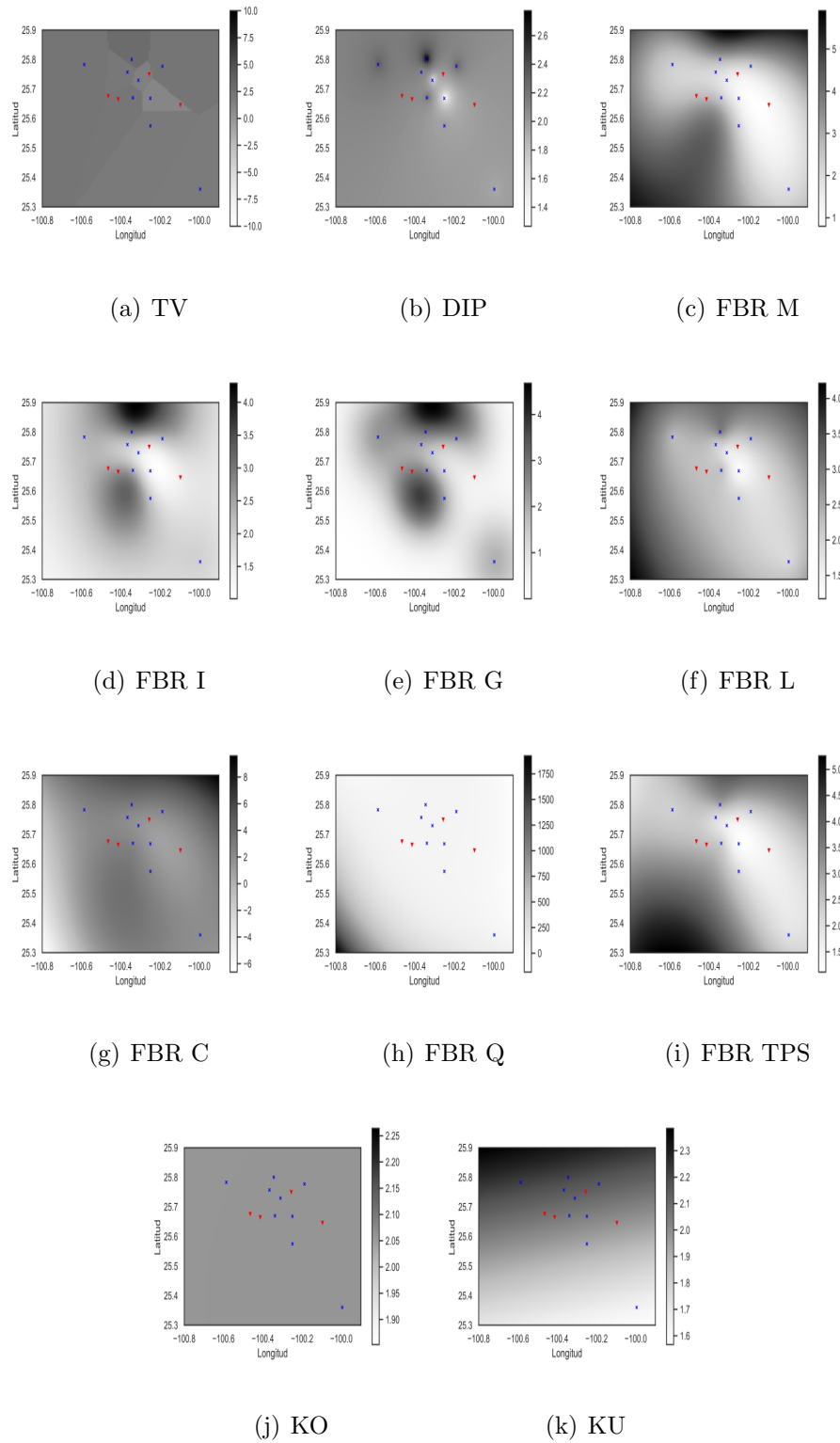


FIGURA 6.21: Interpolaciones de SO_2 para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.22: SO₂: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.72	4.86	0.62	6.78	0.87	45.96
DIP	0.45	3.13	0.40	5.03	0.64	25.37
FBR M	4.36	45.72	5.86	4,732.12	607.40	22,392,980.83
FBR IM	0.94	6.45	0.82	23.81	3.05	567.29
FBR G	28.40	496.80	63.76	133,062.47	1,7079.47	1.77×10^{10}
FBR L	1.87	10.70	1.37	36.29	4.65	1,317.66
FBR C	1.98×10^{13}	7.12×10^{13}	9.14×10^{13}	1.57×10^{16}	2.02×10^{15}	2.48×10^{32}
FBR Q	2.08×10^{17}	1.16×10^{18}	1.50×10^{17}	3.27×10^{18}	4.20×10^{17}	1.07×10^{37}
FBR TPS	1.20	7.42	0.95	24.57	3.15	603.75
KO	0.46	3.18	0.40	5.08	0.65	25.84
KU	0.55	3.72	0.47	6.53	0.83	42.65

De la tabla 6.22, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial (FBR C y FBR Q) de los cuales el mejor prediciendo es el método inverso multicuadrático (FBR IM); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, ya que en todos los errores son menores a los errores de KO. En la figura 6.22, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

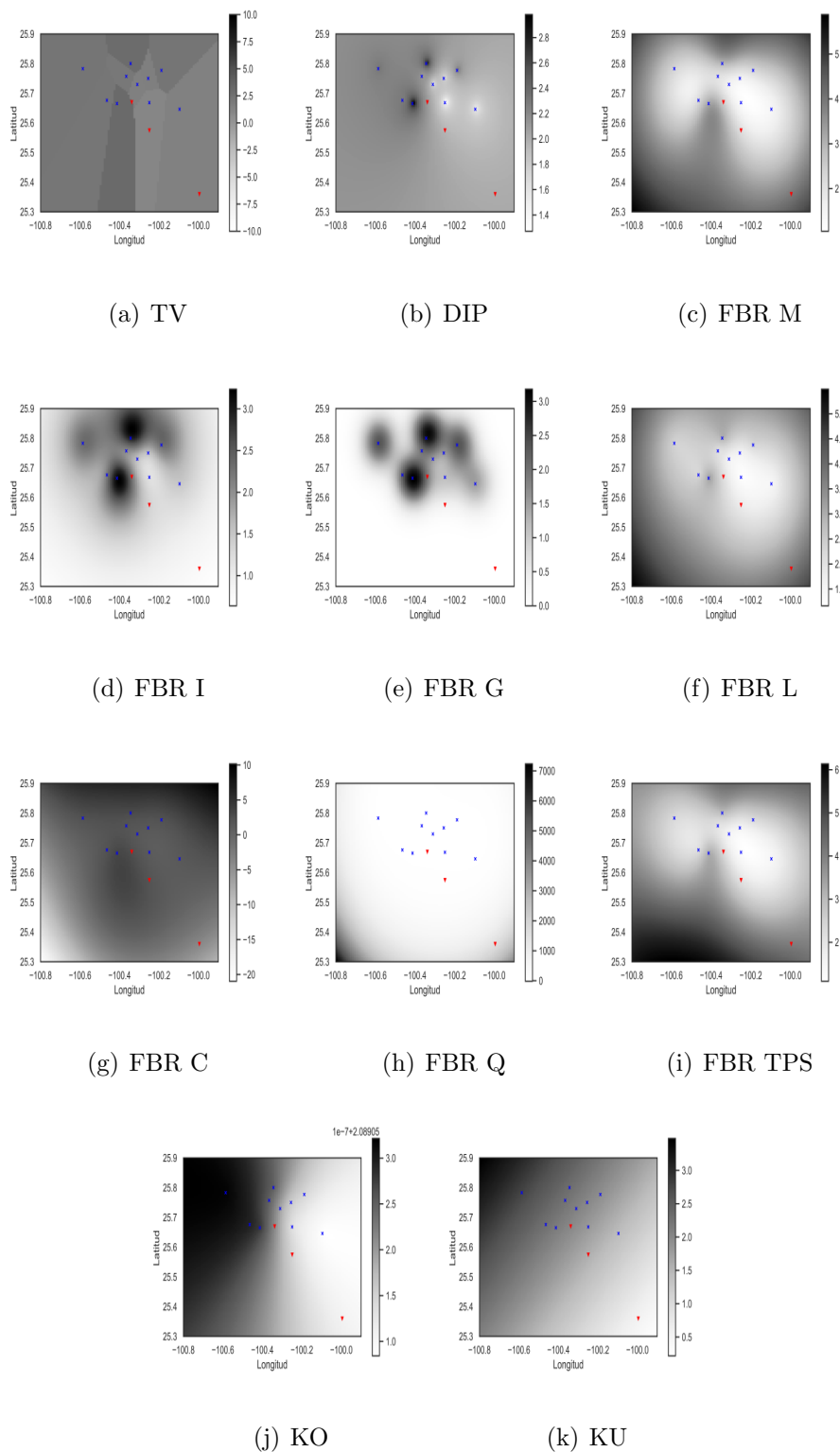


FIGURA 6.22: Interpolaciones de SO_2 para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.23: SO₂: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.73	4.91	0.63	6.76	0.86	45.82
DIP	0.44	3.09	0.39	4.92	0.63	24.24
FBR M	3.54	24.07	3.09	411.24	52.83	169,121.94
FBR IM	1.01	6.76	0.86	18.68	2.40	349.06
FBR G	2.71	18.94	2.43	357.19	45.88	127,587.26
FBR L	1.60	9.35	1.20	34.90	4.48	1,218.50
FBR C	3.03×10^{11}	7.06×10^{12}	9.07×10^{11}	1.61×10^{13}	2.07×10^{14}	2.61×10^{30}
FBR Q	1.66×10^{17}	8.92×10^{17}	1.14×10^{17}	2.86×10^{18}	3.68×10^{17}	8.20×10^{36}
FBR TPS	1.07	6.79	0.87	19.80	2.54	392.42
KO	0.46	3.16	0.40	4.98	0.64	24.83
KU	0.53	3.56	0.45	5.86	0.75	34.42

De la tabla 6.23, en la cual se utilizan once estaciones para interpolar dos más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial (FBR C y FBR Q) de los cuales el mejor prediciendo es el método inverso multicuadrático (FBR IM); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, ya que en todos los errores menores a los errores encontrados por KO. En la figura 6.23, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

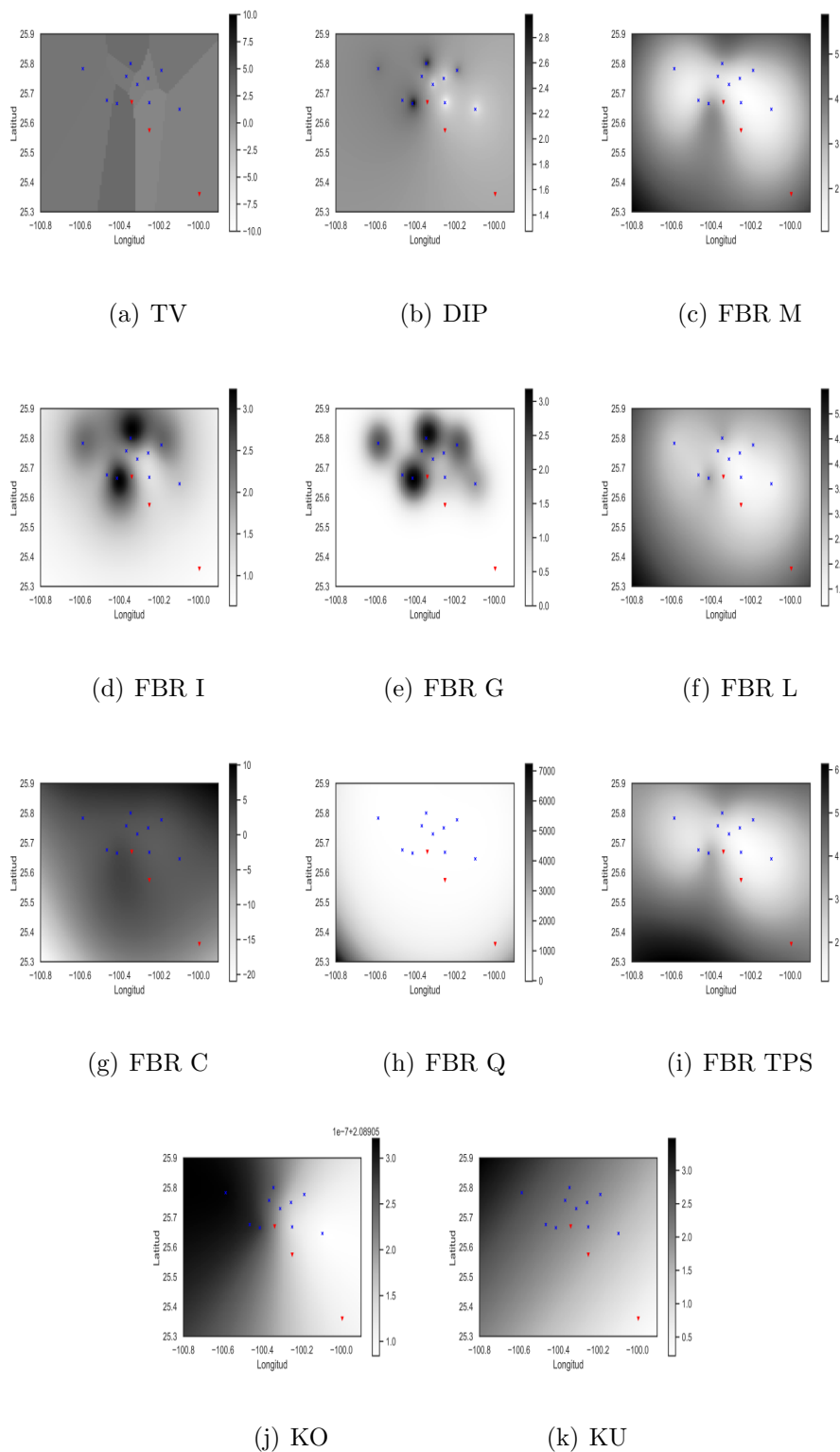


FIGURA 6.23: Interpolaciones de SO_2 para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.24: SO₂: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.73	4.92	0.62	6.80	0.86	46.24
DIP	0.44	3.09	0.39	5.03	0.64	25.33
FBR M	3.14	19.32	2.46	140.57	17.95	19,760.88
FBR IM	1.01	6.59	0.84	12.55	1.60	157.53
FBR G	2.44	14.27	1.82	80.71	10.31	6,514.82
FBR L	1.41	8.28	1.05	21.093	2.69	444.93
FBR C	7.95×10^8	1.06^{10}	1.36×10^9	1.72×10^{12}	2.20×10^{11}	2.98×10^{24}
FBR Q	1.55×10^{17}	8.04×10^{17}	1.02×10^{17}	2.71×10^{18}	3.47×10^{17}	7.39×10^{36}
FBR TPS	1.01	6.41	0.81	11.36	1.45	129.26
KO	0.45	3.16	0.40	5.08	0.64	25.81
KU	0.50	3.48	0.44	5.48	0.70	30.03

De la tabla 6.24, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial (FBR C y FBR Q) de los cuales el mejor prediciendo es el método inverso multicuadrático (FBR IM), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos, KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en todos los errores son menores que los errores de KO. En la figura 6.24, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general mientras se aumenta el número de estaciones para interpolar la variable SO₂, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

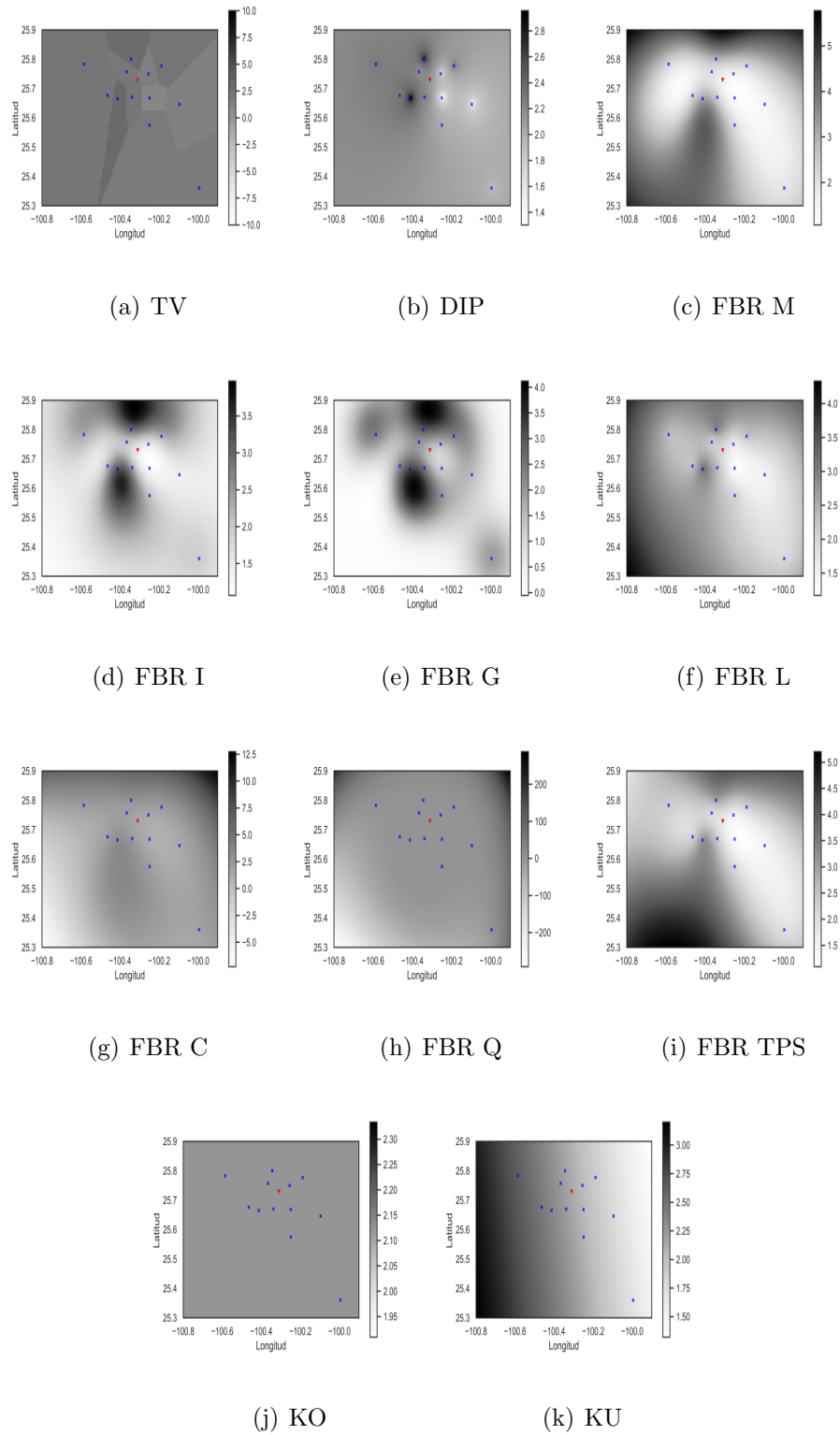


FIGURA 6.24: Interpolaciones de SO_2 para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.2. Resto de Variables

Ahora se presentan los resultados obtenidos de las variables restantes, que nos son consideradas en el *Índice de AIRE y SALUD*, como son: NO, NO_x, *velocidad del viento y dirección del viento*.

6.2.1 ÓXIDO NÍTRICO (NO)

TABLA 6.25: NO: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	2.35	23.65	1.08	54.51	2.49	2,971.71
DIP	1.37	17.77	0.81	47.49	2.17	2,255.77
FBR M	2.55×10^{15}	7.29×10^{16}	3.34×10^{15}	8.19×10^{17}	3.75×10^{16}	6.71×10^{35}
FBR IM	1,940.58	144,806.62	6,635.57	46,942,420.03	2,151,073.61	2.20×10^{15}
FBR G	3.56×10^{13}	9.64×10^{14}	4.41×10^{13}	9.43×10^{16}	4.32×10^{15}	8.89×10^{33}
FBR L	3.65×10^{14}	1.30×10^{16}	5.99×10^{14}	3.44×10^{17}	1.58×10^{16}	1.18×10^{35}
FBR C	3.62×10^{16}	3.29×10^{17}	1.51×10^{16}	1.74×10^{18}	7.98×10^{16}	3.04×10^{36}
FBR Q	2.66×10^{17}	1.74×10^{18}	7.99×10^{16}	4.01×10^{18}	1.83×10^{17}	1.60×10^{37}
FBR TPS	1.02×10^{15}	3.51×10^{16}	161×10^{15}	5.68×10^{17}	2.60×10^{16}	3.23×10^{35}
KO	1.46	17.75	0.81	47.33	2.16	2240.84
KU	23,060.77	1,309,957.63	60,027.05	421,816,508.68	19,329,177.36	1.77×10^{17}

De la tabla 6.25, en la cual se utilizan nueve estaciones para interpolar cuatro estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción del método inverso multicuadrático (FBR IM); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que

el resto de los métodos pero DIP es mejor que KO, pues obtuvo mejores resultados a excepción del método MSE. En la figura 6.25, se pueden observar las interpolaciones de cada método.

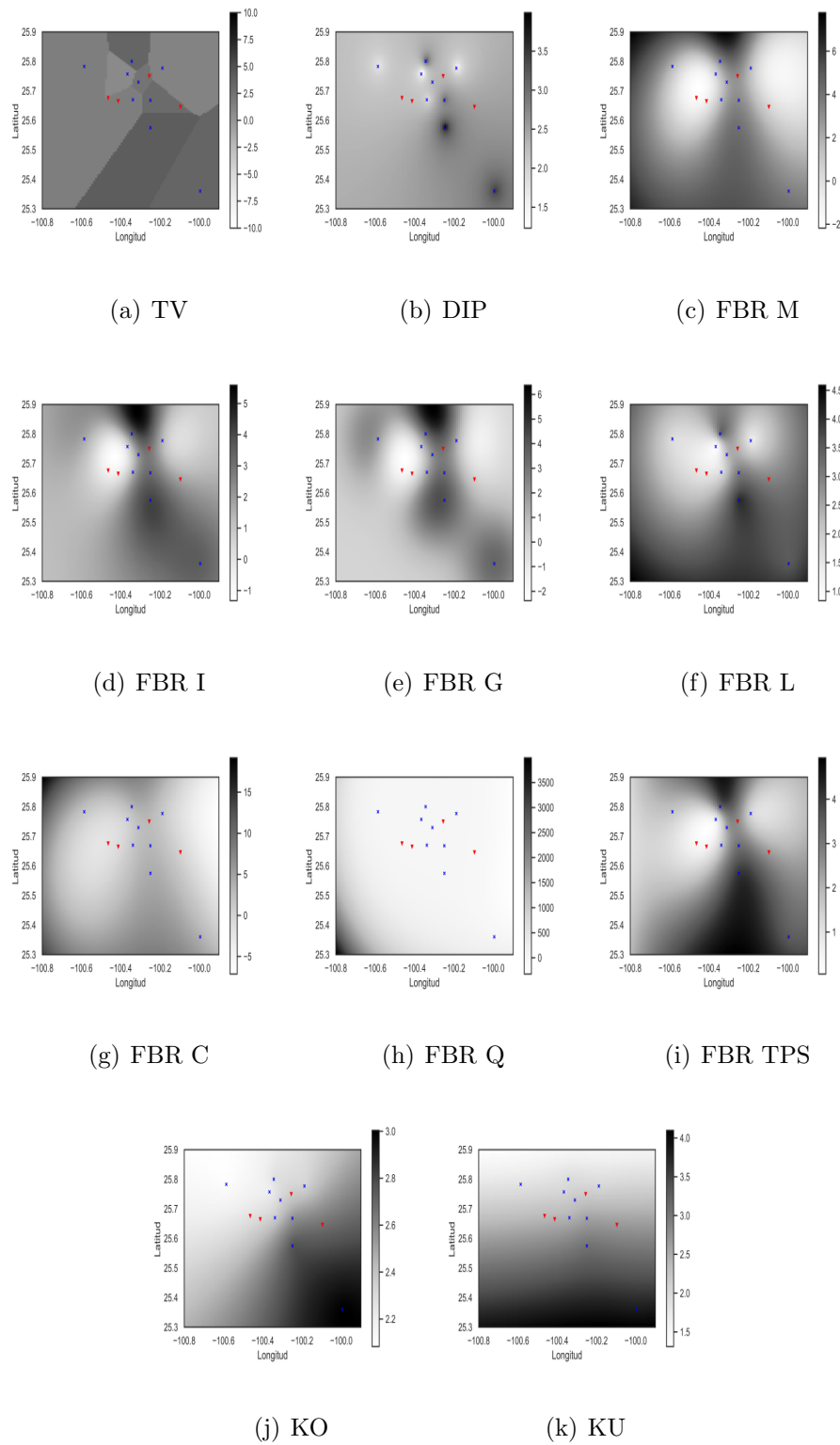


FIGURA 6.25: Interpolaciones de NO para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.26: NO: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	2.40	23.69	1.08	54.3	2.48	2,954.14
DIP	1.36	17.83	0.81	47.86	2.18	2,291.14
FBR M	2.92×10^{15}	8.61×10^{16}	3.94×10^{15}	8.89×10^{17}	4.06×10^{16}	7.91×10^{35}
FBR IM	54.25	2,126.17	97.25	524,486.19	23,990.76	2.75×10^{11}
FBR G	1.13×10^{14}	3.81×10^{15}	1.74×10^{14}	1.86×10^{17}	8.53×10^{15}	3.47×10^{34}
FBR L	1.69×10^{14}	8.84×10^{15}	4.04×10^{14}	2.83×10^{17}	1.29×10^{16}	8.03×10^{34}
FBR C	4.11×10^{16}	3.64×10^{17}	1.66×10^{16}	1.83×10^{18}	8.38×10^{16}	3.36×10^{36}
FBR Q	2.93×10^{17}	1.97×10^{18}	9.02×10^{16}	4.26×10^{18}	1.95×10^{17}	1.81×10^{37}
FBR TPS	7.06×10^{14}	3.09×10^{16}	1.41×10^{15}	5.32×10^{17}	2.43×10^{16}	2.83×10^{35}
KO	1.43	17.74	0.81	47.61	2.17	2,267.56
KU	7.40	90.72	4.14	4,855.33	222.09	2.35×10^7

De la tabla 6.26, en la cual se utilizan diez estaciones para interpolar tres más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción del método inverso multicuadrático (FBR IM), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. En general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son mejores que los resultados de KO a excepción del método MSE, donde KO obtuvo un mejor resultado que DIP. En la figura 6.26, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

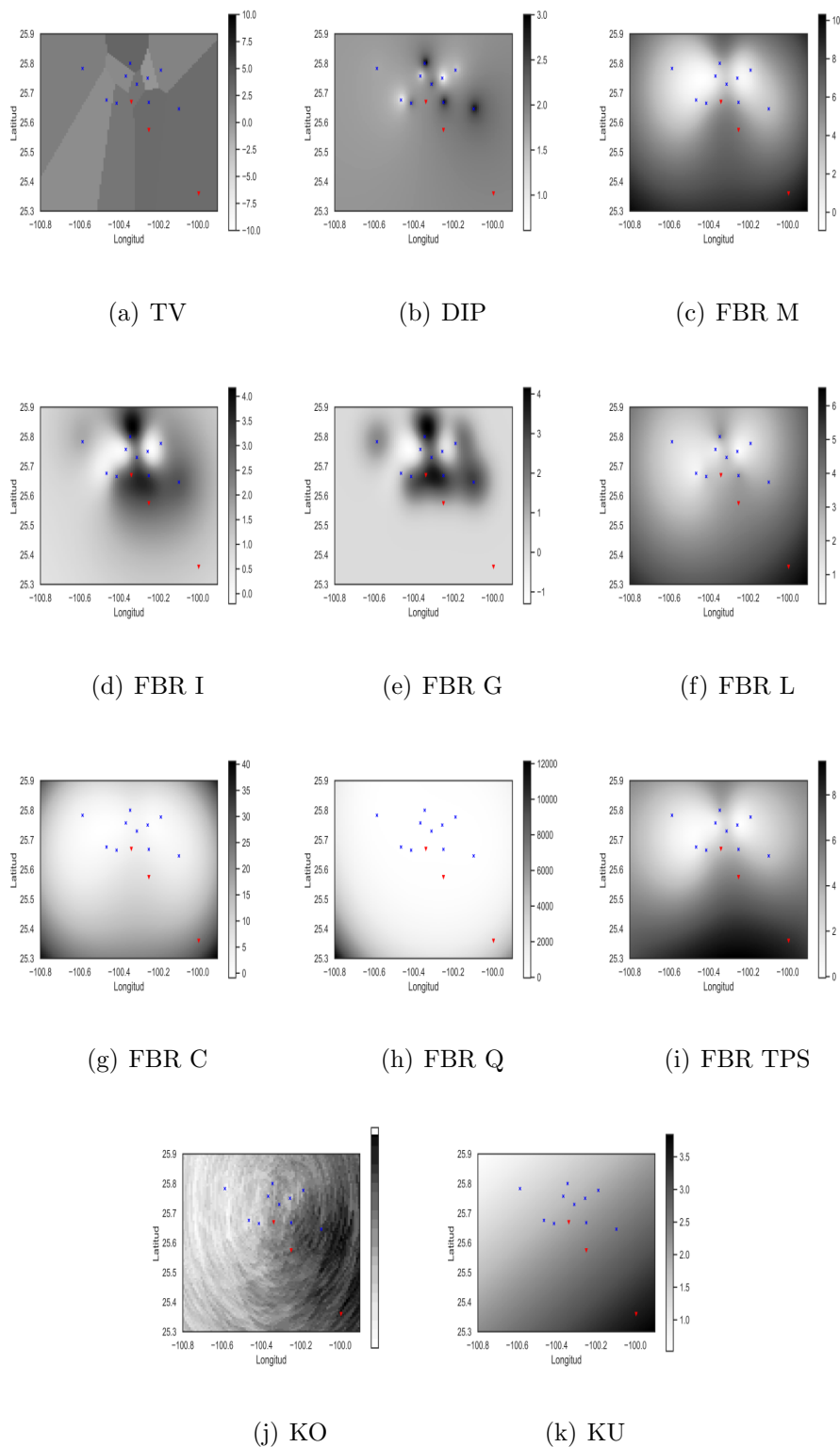


FIGURA 6.26: Interpolaciones de NO para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.27: NO: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	2.38	23.91	1.08	54.36	2.45	2,955.92
DIP	1.33	18.04	0.81	48.49	2.19	2,351.50
FBR M	3.31×10^{15}	1.04×10^{17}	4.72×10^{15}	9.80×10^{17}	4.43×10^{16}	9.61×10^{35}
FBR IM	7.90	85.75	3.87	1,800.30	81.371	3,241,097.69
FBR G	3.59×10^{14}	9.75×10^{15}	4.40×10^{14}	2.98×10^{17}	1.34×10^{16}	8.90×10^{34}
FBR L	7.86×10^{14}	4.21×10^{15}	1.90×10^{14}	1.97×10^{17}	8.90×10^{15}	3.88×10^{34}
FBR C	3.86×10^{16}	3.57×10^{17}	1.61×10^{16}	1.81×10^{18}	8.20×10^{16}	3.29×10^{36}
FBR Q	2.37×10^{17}	1.67×10^{18}	7.58×10^{16}	3.93×10^{18}	1.77×10^{17}	1.54×10^{37}
FBR TPS	7.54×10^{14}	3.70×10^{16}	1.67×10^{15}	5.82×10^{17}	2.63×10^{16}	3.39×10^{35}
KO	1.40	17.93	0.81	48.23	2.18	2,326.83
KU	3.12	40.44	1.82	1,148.871	51.92	1,319,905.49

De la tabla 6.27, en la cual se utilizan once estaciones para interpolar dos más, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción del método inverso multicuadrático (FBR IM), entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son menores o iguales a los errores de KO, a excepción del método MSE. En la figura 6.27, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

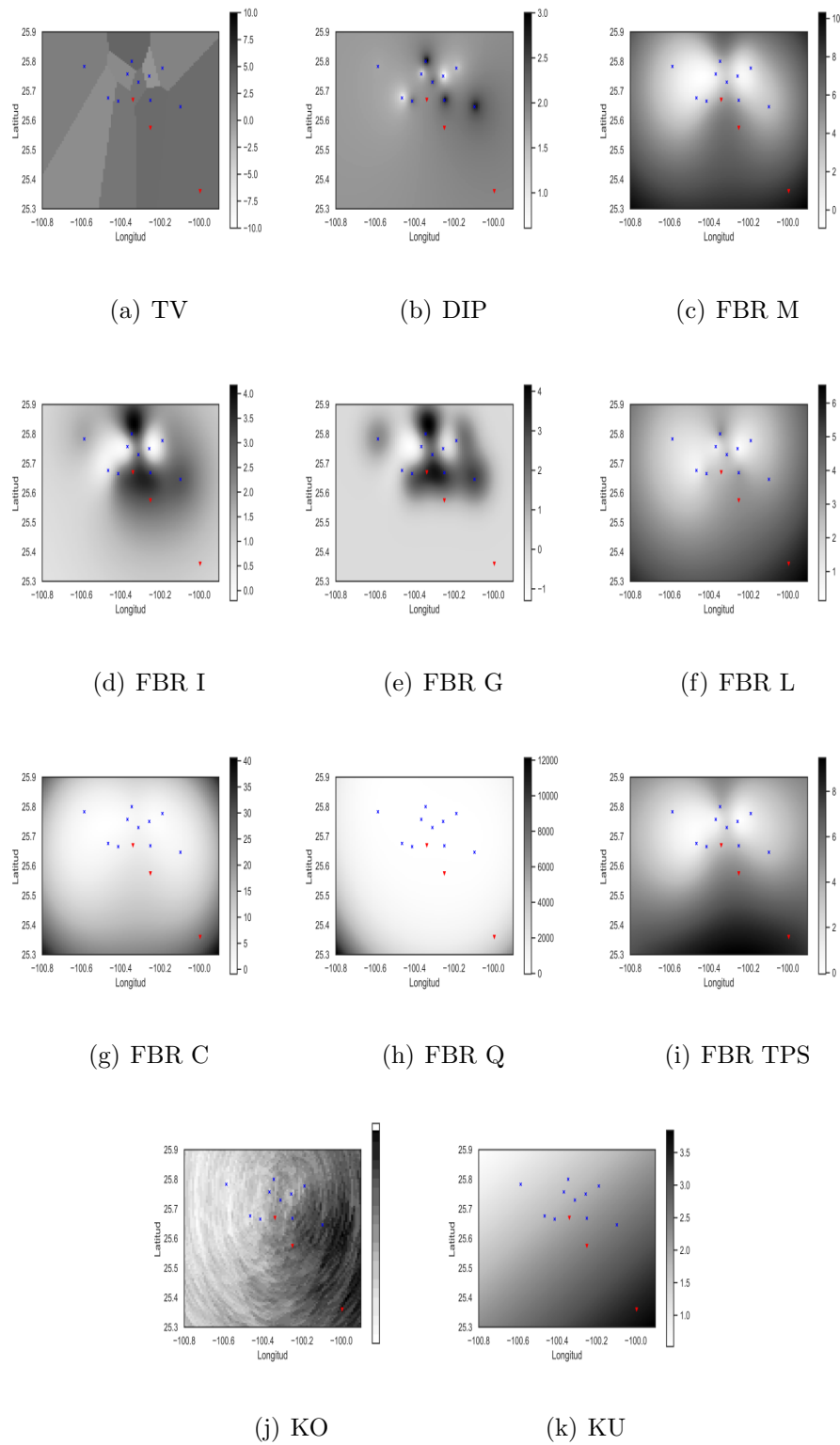


FIGURA 6.27: Interpolaciones de NO para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.28: NO: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	2.40	23.44	1.08	52.81	2.44	2,789.82
DIP	1.35	17.51	0.81	47.40	2.19	2,247.21
FBR M	3.14×10^{15}	1.22×10^{17}	5.69×10^{15}	1.06×10^{18}	4.92×10^{16}	1.12×10^{36}
FBR IM	9.11	110.26	5.11	1,690.80	78.40	2,858,836.55
FBR G	5.79×10^{15}	2.00×10^{16}	9.30×10^{14}	4.29×10^{17}	1.99×10^{16}	1.84×10^{35}
FBR L	8,271.3	436,546.93	20,242.38	7,043,358.60	326,595.78	4.96×10^{13}
FBR C	3.34×10^{16}	3.19×10^{17}	1.48×10^{16}	1.71×10^{18}	7.95×10^{16}	2.94×10^{36}
FBR Q	1.50×10^{17}	1.18×10^{18}	5.48×10^{16}	3.30×10^{18}	1.53×10^{17}	1.09×10^{37}
FBR TPS	1.24×10^{16}	6.06×10^{16}	2.81×10^{15}	7.45×10^{17}	3.45×10^{16}	5.55×10^{35}
KO	1.42	17.39	0.80	47.11	2.18	2,219.88
KU	2.44	25.74	1.19	100.10	4.64	10,021.27

De la tabla 6.28, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, a excepción de los métodos lineal e inverso multicuadrático (FBR L y FBR IM); entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son mejores que los errores de KO, a excepción del método MSE. En la figura 6.28 se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general aumenta el número de estaciones para interpolar la variable NO, estas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

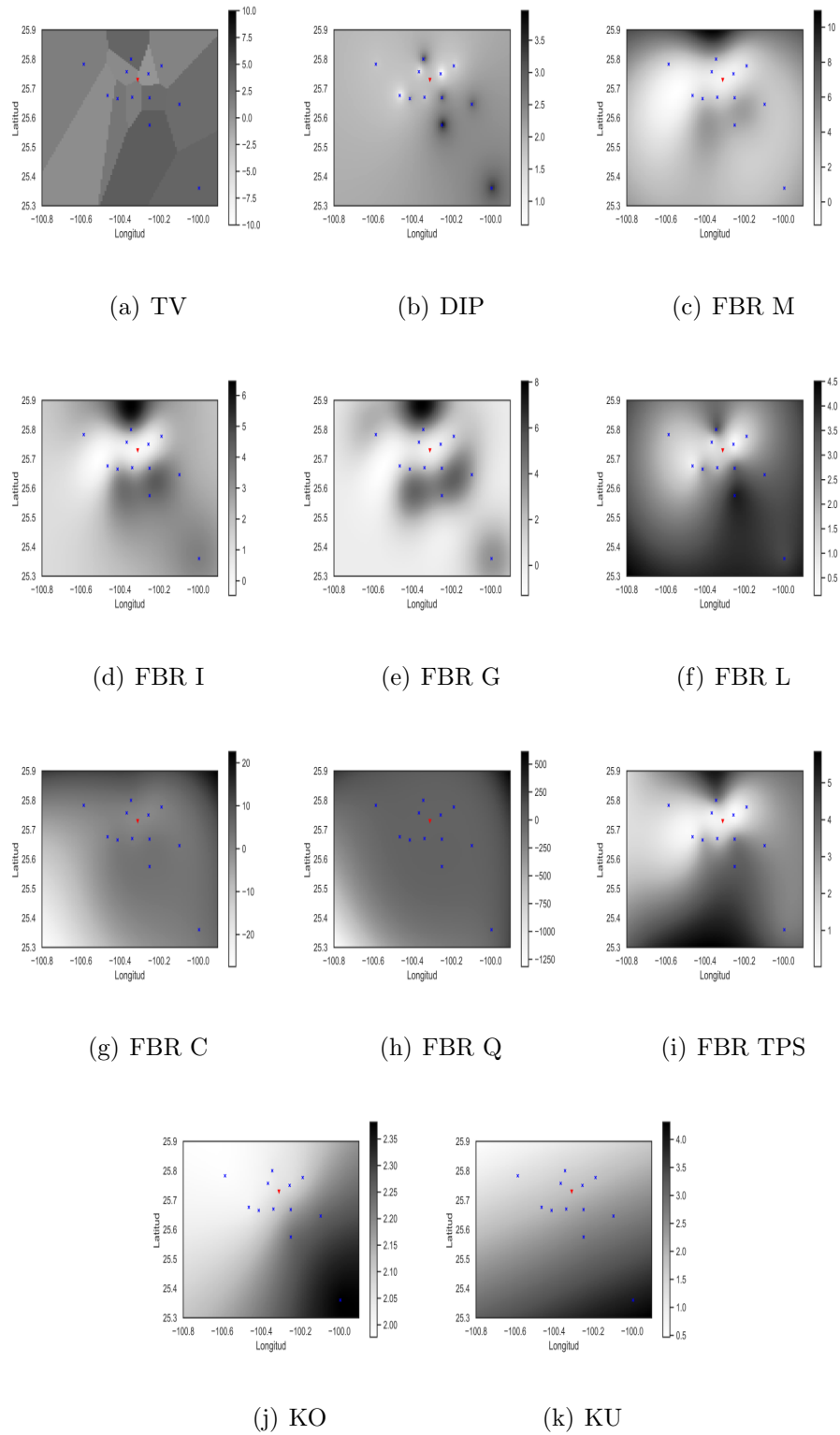


FIGURA 6.28: Interpolaciones de NO para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.2.2 ÓXIDOS DE NITRÓGENO (NO_X)TABLA 6.29: NO_X : 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.69	40.01	1.01	72.86	1.85	5,308.96
DIP	0.79	25.97	0.66	54.97	1.40	3,022.32
FBR M	2.39×10^{16}	4.92×10^{17}	1.25×10^{16}	2.13×10^{18}	5.43×10^{16}	4.54×10^{36}
FBR IM	1.36×10^{13}	5.25×10^{14}	1.34×10^{13}	6.96×10^{16}	1.77×10^{15}	4.85×10^{33}
FBR G	4.47×10^{14}	7.71×10^{15}	1.96×10^{14}	2.66×10^{17}	6.79×10^{15}	7.11×10^{34}
FBR L	3.97×10^{16}	6.55×10^{17}	1.66×10^{16}	2.45×10^{18}	6.26×10^{16}	6.04×10^{36}
FBR C	8.69×10^{15}	2.19×10^{17}	5.59×10^{15}	1.42×10^{18}	3.62×10^{16}	2.02×10^{36}
FBR Q	1.33×10^{17}	2.32×10^{18}	5.91×10^{16}	4.62×10^{18}	1.17×10^{17}	2.14×10^{37}
FBR TPS	1.38×10^{16}	3.18×10^{17}	8.11×10^{15}	1.71×10^{18}	4.36×10^{16}	2.93×10^{36}
KO	0.85	26.46	0.67	55.40	1.41	3,070.18
KU	3.00×10^{13}	1.13×10^{15}	2.90×10^{13}	1.02×10^{17}	2.61×10^{15}	1.05×10^{34}

De la tabla 6.29, en la cual se utilizan nueve estaciones para interpolar cuatro estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son más pequeños los errores de KO. En la figura 6.29, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

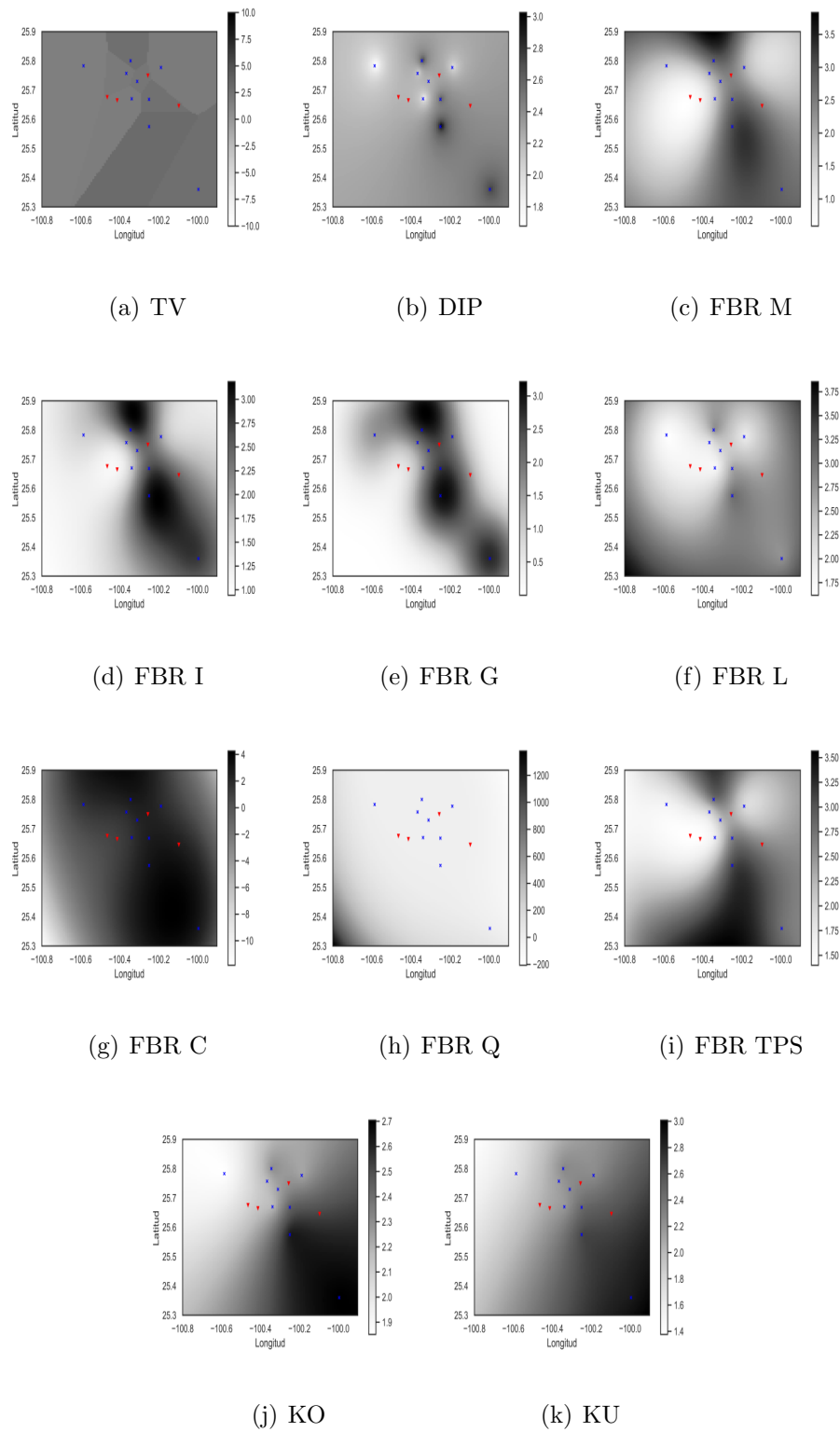


FIGURA 6.29: Interpolaciones de NO_x para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.30: NO_x : 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.78	40.85	1.04	73.67	1.88	5,427.87
DIP	0.78	25.89	0.66	55.15	1.41	3,041.99
FBR M	2.23×10^{16}	4.73×10^{17}	1.21×10^{16}	2.09×10^{18}	5.34×10^{16}	4.37×10^{36}
FBR IM	1.08×10^{14}	2.33×10^{14}	5.97×10^{12}	4.64×10^{16}	1.18×10^{15}	2.15×10^{33}
FBR G	4.79×10^{14}	8.36×10^{15}	2.13×10^{14}	2.77×10^{17}	7.09×10^{15}	7.68×10^{34}
FBR L	3.35×10^{16}	5.80×10^{17}	1.48×10^{16}	2.31×10^{18}	5.91×10^{16}	5.35×10^{36}
FBR C	8.60×10^{15}	2.08×10^{17}	5.34×10^{12}	1.38×10^{18}	3.55×10^{16}	1.92×10^{36}
FBR Q	1.48×10^{17}	2.51×10^{18}	6.42×10^{16}	4.81×10^{18}	1.23×10^{17}	2.31×10^{37}
FBR TPS	1.00×10^{16}	2.55×10^{17}	6.52×10^{15}	1.53×10^{18}	3.92×10^{16}	2.35×10^{36}
KO	0.84	26.28	0.67	55.19	1.41	3,045.95
KU	1.66×10^{13}	4.67×10^{14}	1.19×10^{13}	6.56×10^{16}	1.67×10^{15}	4.31×10^{33}

De la tabla 6.30, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores, éstos suelen ser menores a los errores de KO. En la figura 6.30, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

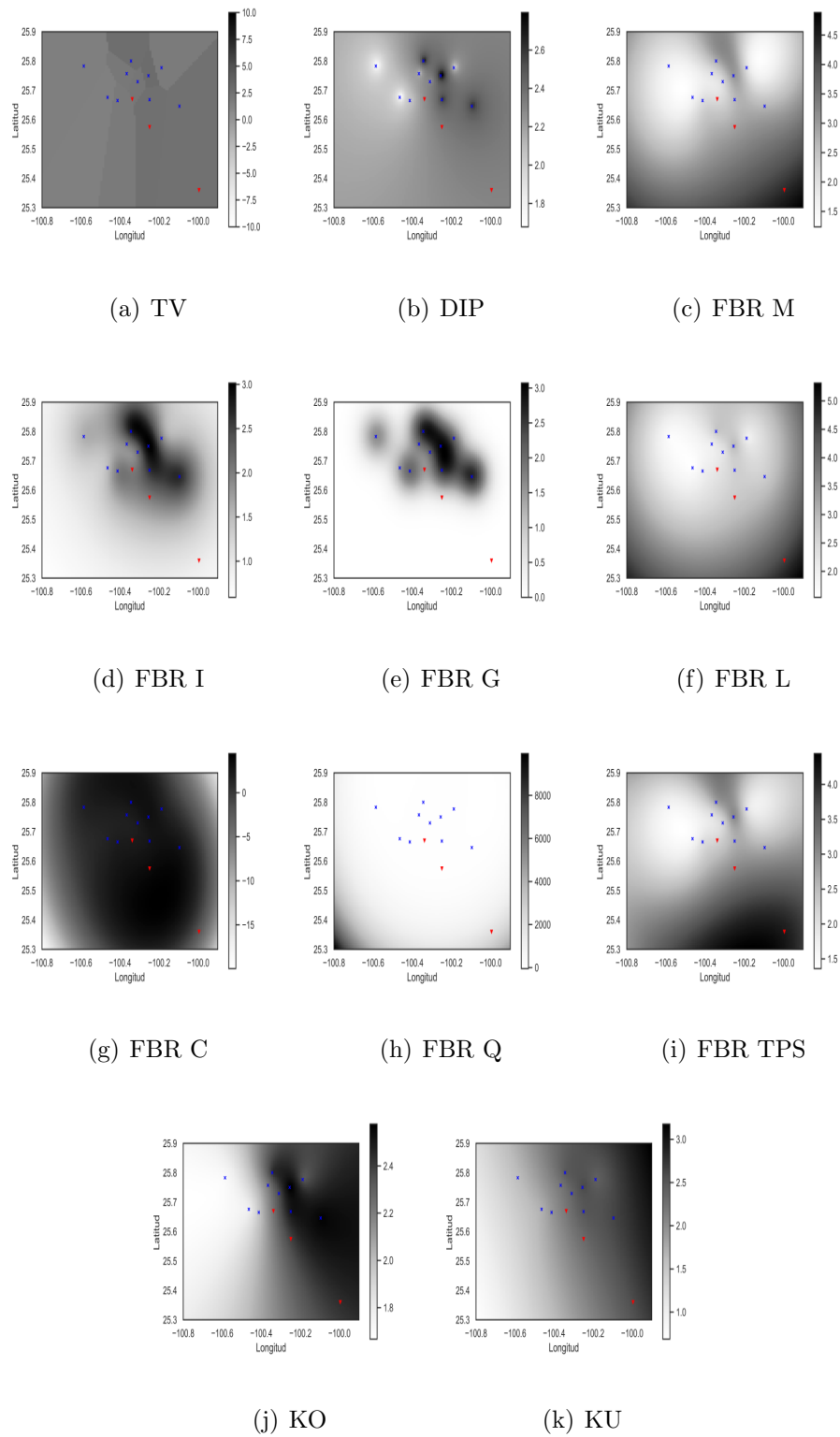


FIGURA 6.30: Interpolaciones de NO_x para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.31: NO_x: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.87	41.94	1.06	74.88	1.90	5,607.14
DIP	0.78	25.99	0.66	55.52	1.41	3,082.71
FBR M	2.22×10^{16}	4.78×10^{17}	1.21×10^{16}	2.09×10^{18}	5.34×10^{16}	4.40×10^{36}
FBR IM	7.87×10^{12}	5.25×10^{14}	1.33×10^{13}	6.96×10^{16}	1.77×10^{15}	4.85×10^{33}
FBR G	8.14×10^{14}	1.85×10^{16}	4.72×10^{14}	4.14×10^{17}	1.05×10^{16}	1.71×10^{35}
FBR L	2.81×10^{16}	5.18×10^{17}	1.31×10^{16}	2.18×10^{18}	5.56×10^{16}	4.78×10^{36}
FBR C	9.70×10^{15}	2.39×10^{17}	6.08×10^{15}	1.48×10^{18}	3.77×10^{16}	2.20×10^{36}
FBR Q	1.17×10^{17}	1.96×10^{18}	5.00×10^{16}	4.26×10^{18}	1.08×10^{17}	1.81×10^{37}
FBR TPS	7.42×10^{15}	2.26×10^{17}	5.76×10^{16}	1.44×10^{18}	3.67×10^{16}	2.08×10^{36}
KO	0.82	26.24	0.66	55.30	1.40	3,059.11
KU	9.11×10^{12}	3.50×10^{14}	8.92×10^{12}	5.68×10^{16}	1.44×10^{15}	3.23×10^{33}

De la tabla 6.31, en la cual se utilizan once estaciones para interpolar dos estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son menores que los errores de KO, a excepción del método MSE. En la figura 6.31, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

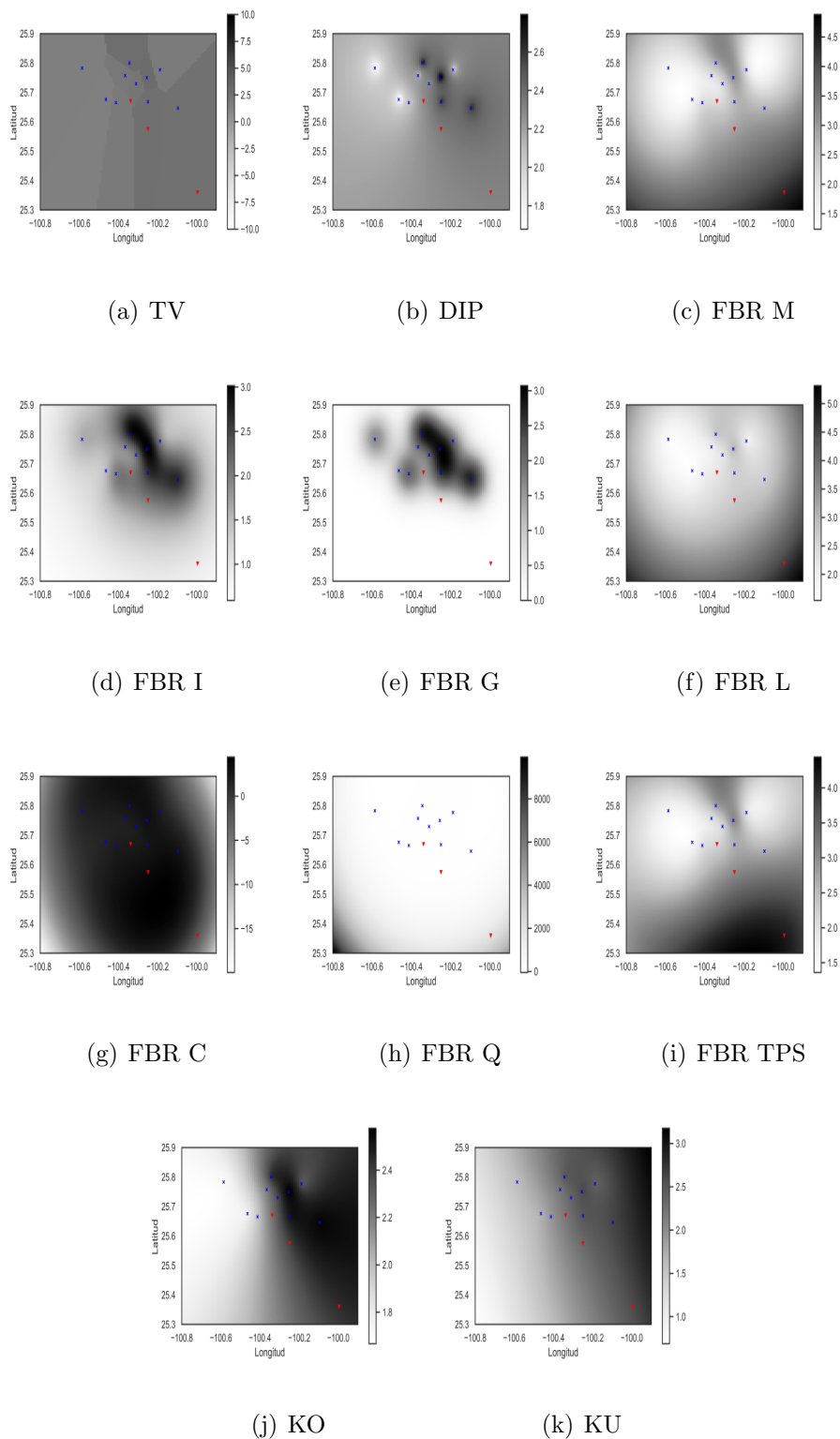


FIGURA 6.31: Interpolaciones de NO_x para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.32: NO_X : 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	1.90	42.12	1.07	74.58	1.90	5,562.30
DIP	0.77	25.71	0.65	54.91	1.40	3,015.47
FBR M	2.43×10^{16}	4.96×10^{17}	1.27×10^{16}	2.14×10^{18}	5.47×10^{16}	4.58×10^{36}
FBR IM	34.52	475.91	12.16	58,490.29	1,495.41	3,421,114,460.30
FBR G	8.44×10^{14}	2.62×10^{16}	6.72×10^{14}	4.92×10^{17}	1.25×10^{16}	2.42×10^{35}
FBR L	2.61×10^{16}	5.00^{17}	1.27×10^{16}	2.14×10^{18}	5.49×10^{16}	4.61×10^{36}
FBR C	1.06×10^{16}	2.35×10^{17}	6.01×10^{15}	1.47×10^{18}	3.76×10^{16}	2.17×10^{36}
FBR Q	8.22×10^{16}	1.40×10^{18}	3.60×10^{16}	3.60×10^{18}	9.21×10^{16}	1.30×10^{37}
FBR TPS	4.56×10^{15}	1.87×10^{17}	4.80×10^{15}	1.31×10^{18}	3.36×10^{16}	1.73×10^{36}
KO	0.81	25.90	0.66	54.61	1.39	2,982.28
KU	1.24	35.30	0.90	119.61	3.05	14,308.07

De la tabla 6.32, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE. Los métodos DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son menores a los errores de KO, a excepción del método MSE. En la figura 6.32, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general mientras se aumenta el número de estaciones para interpolar la variable NO_X , éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

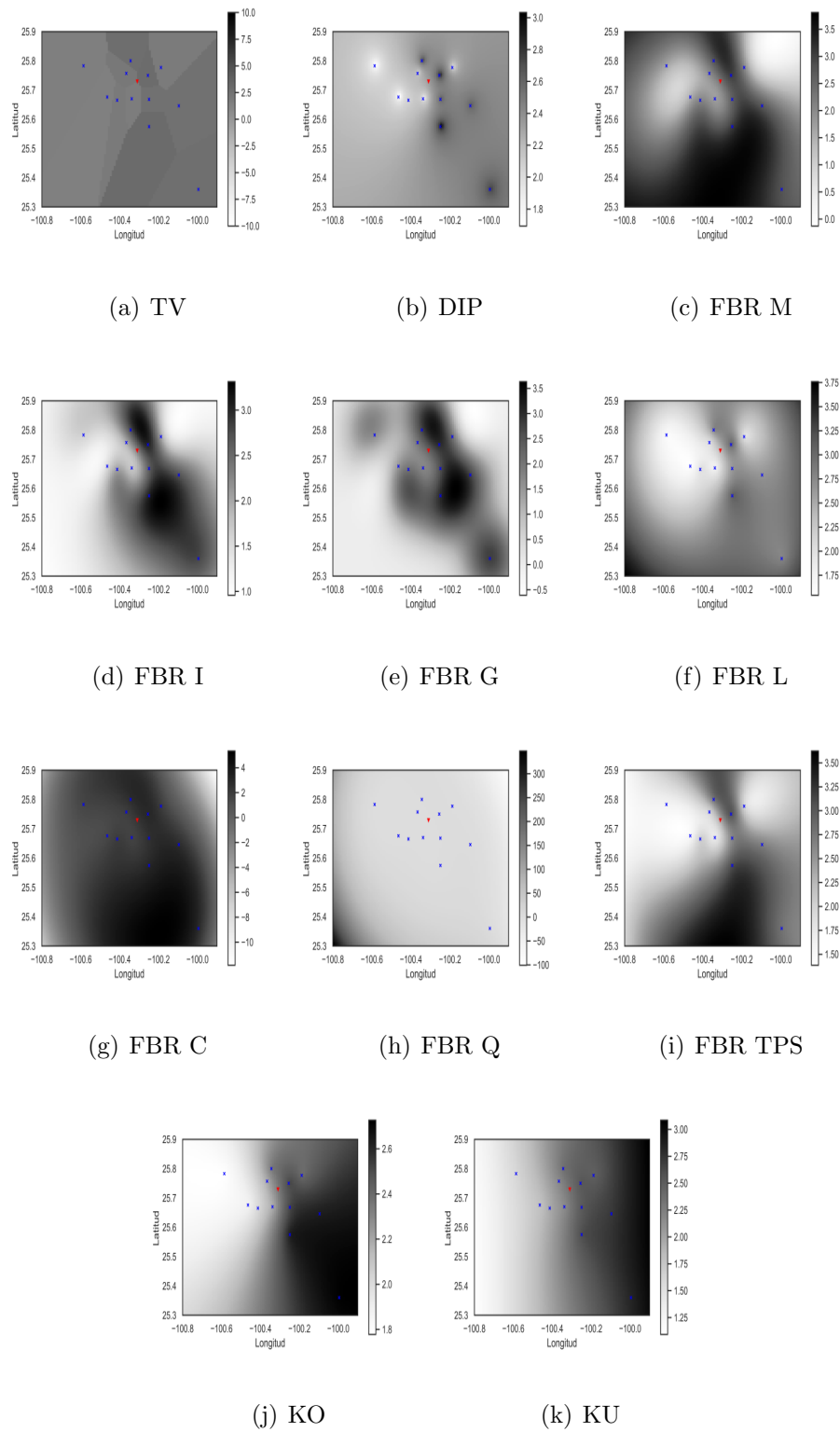


FIGURA 6.32: Interpolaciones de NO_x para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.2.3 VELOCIDAD DEL VIENTO

TABLA 6.33: *Velocidad del viento*: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.76	3.76	0.49	5.084	0.66	25.85
DIP	0.53	2.68	0.35	3.79	0.49	14.42
FBR M	1.63×10^{16}	5.67×10^{16}	7.43×10^{15}	7.23×10^{17}	9.47×10^{16}	5.23×10^{35}
FBR IM	8.81×10^{14}	3.06×10^{13}	4.02×10^{14}	1.68×10^{17}	2.20×10^{16}	2.82×10^{34}
FBR G	1.47×10^{16}	5.61×10^{16}	7.35×10^{15}	7.19×10^{17}	9.42×10^{16}	5.17×10^{35}
FBR L	1.31	6.50	0.85	14.66	1.92	214.98
FBR C	1.99×10^{17}	7.89×10^{17}	1.03×10^{17}	2.69×10^{18}	3.53×10^{17}	7.27×10^{36}
FBR Q	4.51×10^{17}	1.93×10^{18}	2.53×10^{17}	4.22×10^{18}	5.53×10^{17}	1.78×10^{37}
FBR TPS	1.48×10^{13}	8.76×10^{13}	1.14×10^{12}	2.84×10^{16}	3.72×10^{15}	8.08×10^{32}
KO	0.56	2.77	0.36	3.96	0.51	15.72
KU	2.12×10^{14}	4.38×10^{14}	5.74×10^{13}	6.35×10^{16}	8.33×10^{15}	4.04×10^{33}

De la tabla 6.33, en la cual se utilizan nueve estaciones para interpolar otras cuatro, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de funciones de base radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en el resto de los errores con menores que los errores encontrados por KO. En la figura 6.37, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

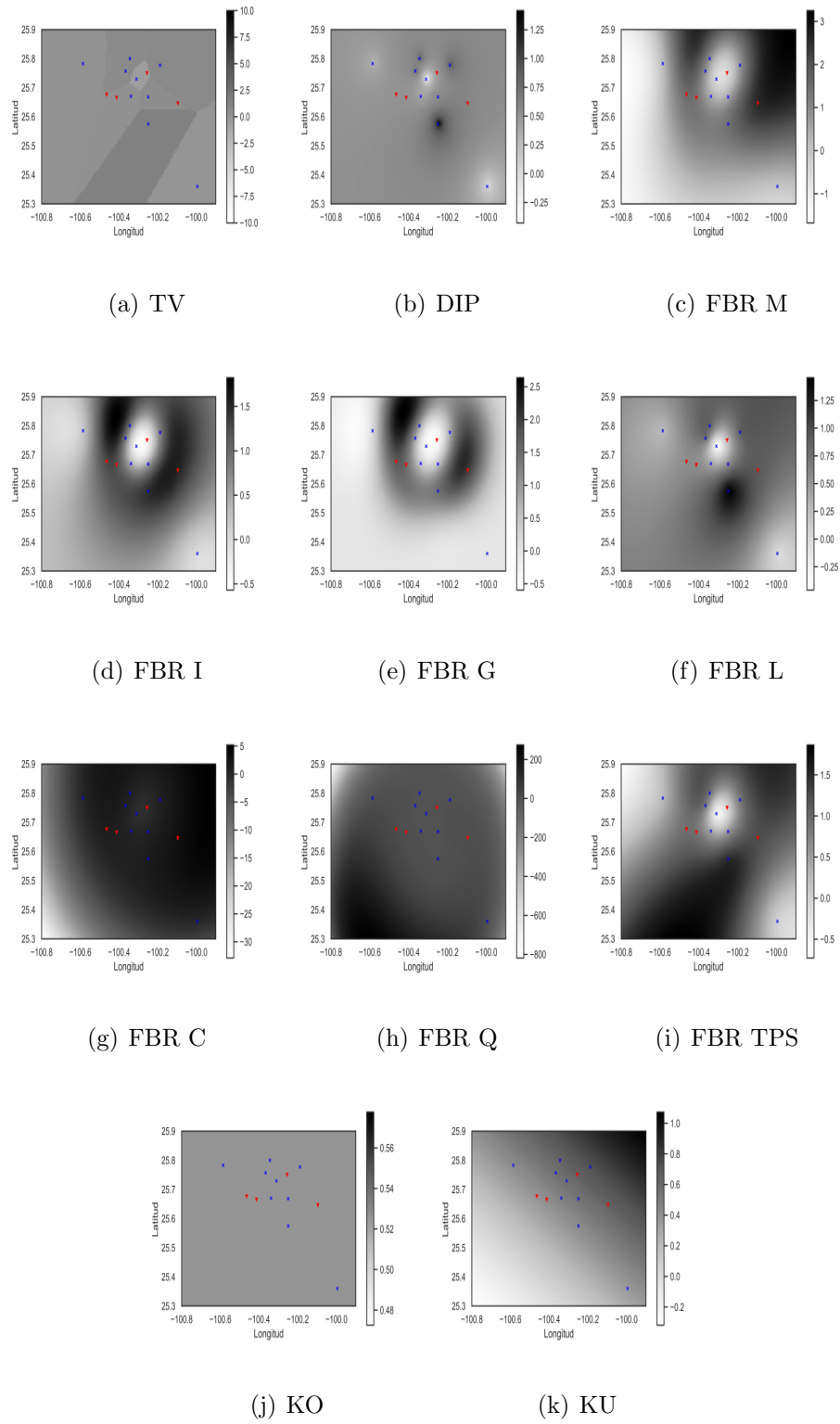


FIGURA 6.33: Interpolaciones de *velocidad del viento* para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.34: *Velocidad del viento*: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.77	3.79	0.49	5.10	0.67	26.09
DIP	0.53	2.66	0.35	3.73	0.49	13.93
FBR M	1.77×10^{16}	7.05×10^{16}	9.29×10^{15}	8.06×10^{17}	1.06×10^{15}	6.51×10^{35}
FBR IM	9.03×10^{14}	3.62×10^{15}	4.77×10^{14}	1.82×10^{17}	2.40×10^{16}	3.34×10^{34}
FBR G	1.64×10^{16}	7.51×10^{16}	9.89×10^{16}	8.32×10^{17}	1.09×10^{17}	6.93×10^{35}
FBR L	1.27	6.19	0.81	13.42	1.76	180.30
FBR C	1.81×10^{17}	7.01×10^{17}	9.23×10^{16}	2.54×10^{18}	3.34×10^{17}	6.46×10^{36}
FBR Q	4.33×10^{17}	1.80×10^{18}	2.38×10^{17}	4.08×10^{18}	5.37×10^{17}	1.66×10^{37}
FBR TPS	1.16	5.61	0.73	10.84	1.42	117.62
KO	0.56	2.76	0.36	3.92	0.51	15.43
KU	0.77	3.57	0.47	5.62	0.73	31.58

De la tabla 6.34, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en el resto de los errores son más pequeños que los errores de KO. En la figura 6.38, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

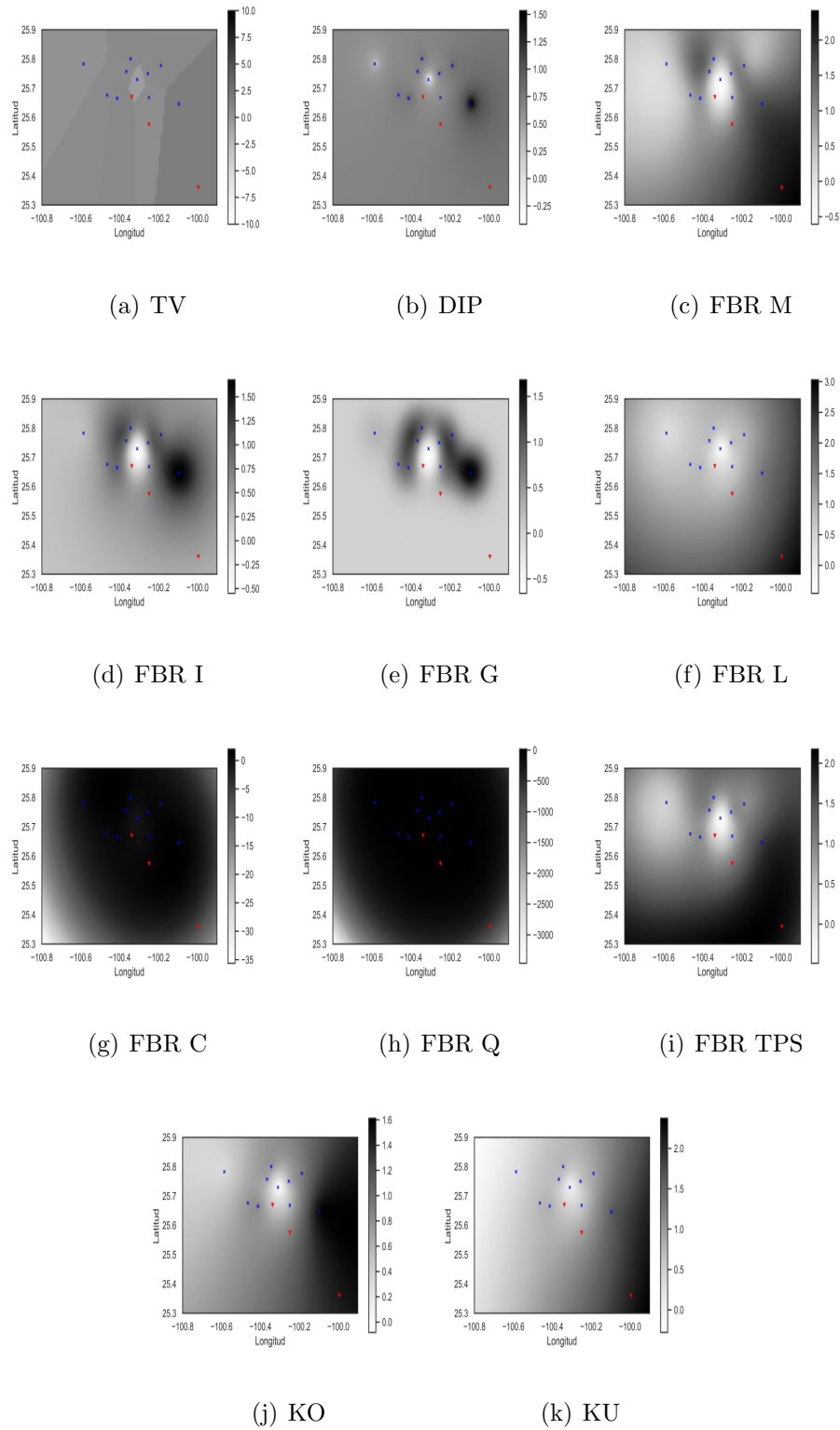


FIGURA 6.34: Interpolaciones de *velocidad del viento* para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.35: *Velocidad del viento*: 11 estaciones seleccionadas 2 estaciones interpoladas

Método	Métricas de error					
	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.78	3.81	0.50	5.12	0.67	26.25
DIP	0.53	2.65	0.34	3.73	0.49	13.96
FBR M	1.65×10^{16}	8.60×10^{16}	1.13×10^{16}	8.91×10^{17}	1.17×10^{17}	7.93×10^{35}
FBR IM	2.26×10^{14}	1.57×10^{15}	2.07×10^{14}	1.20×10^{17}	1.58×10^{16}	1.45×10^{34}
FBR G	1.91×10^{16}	1.10×10^{17}	1.45×10^{16}	1.00×10^{18}	1.32×10^{17}	1.01×10^{36}
FBR L	1.23	6.09	0.80	12.73	1.67	162.11
FBR C	1.59×10^{17}	6.31×10^{17}	8.32×10^{16}	2.41×10^{18}	3.17×10^{17}	5.82×10^{36}
FBR Q	3.24×10^{17}	1.42×10^{18}	1.88×10^{17}	3.63×10^{18}	4.78×10^{17}	1.31×10^{37}
FBR TPS	1.09×10^{14}	1.75×10^{14}	2.30×10^{14}	4.02×10^{16}	5.29×10^{15}	1.61×10^{33}
KO	0.56	2.75	0.36	3.93	0.51	15.46
KU	0.77	3.54	0.46	5.55	0.73	30.80

De la tabla 6.35, en la cual se utilizan once estaciones para interpolar dos estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos, pero DIP es mejor que KO. En la figura 6.39, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

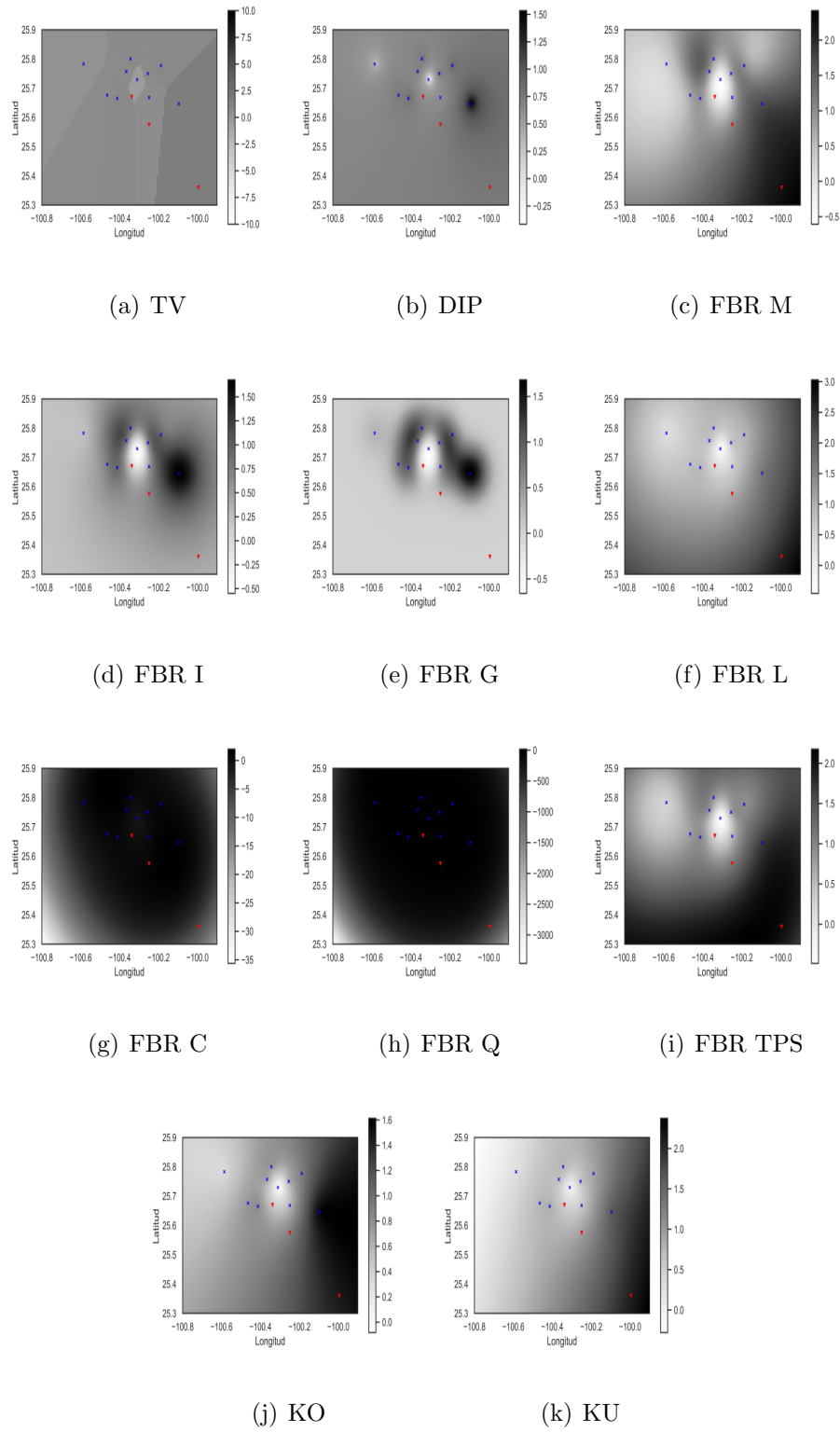


FIGURA 6.35: Interpolaciones de *velocidad del viento* para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.36: *Velocidad del viento*: 12 estaciones seleccionadas 1 estación interpolada

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	0.79	3.83	0.50	5.13	0.67	26.40
DIP	0.52	2.62	0.34	3.69	0.48	13.64
FBR M	1.83×10^{16}	1.35×10^{17}	1.78×10^{16}	1.11×10^{18}	1.46×10^{17}	1.24×10^{36}
FBR IM	0.93	5.09	0.67	7.20	0.94	51.95
FBR G	2.79×10^{16}	2.33×10^{17}	3.07×10^{16}	1.46×10^{18}	1.93×10^{17}	2.15×10^{36}
FBR L	1.21	6.00	0.78	12.21	1.60	149.13
FBR C	1.45×10^{17}	5.38×10^{17}	7.08×10^{16}	2.22×10^{18}	2.93×10^{17}	4.96×10^{36}
FBR Q	1.84×10^{17}	8.86×10^{17}	1.16×10^{17}	2.85×10^{18}	3.76×10^{17}	8.17×10^{36}
FBR TPS	1.06	5.23	0.68	8.22	1.08	67.59
KO	0.56	2.74	0.36	3.91	0.51	15.35
KU	0.76	3.50	0.46	5.48	0.72	30.13

De la tabla 6.36, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en el resto de los errores son menores los errores de KO. En la figura 6.40, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas. En general mientras se aumenta el número de estaciones para interpolar la variable *velocidad del viento*, éstas bajan de forma rápida sus errores en algunos métodos como las funciones de base radial y para el resto de los métodos que ya son buenos para interpolar también mejoran pero los cambios ya son con menor intensidad.

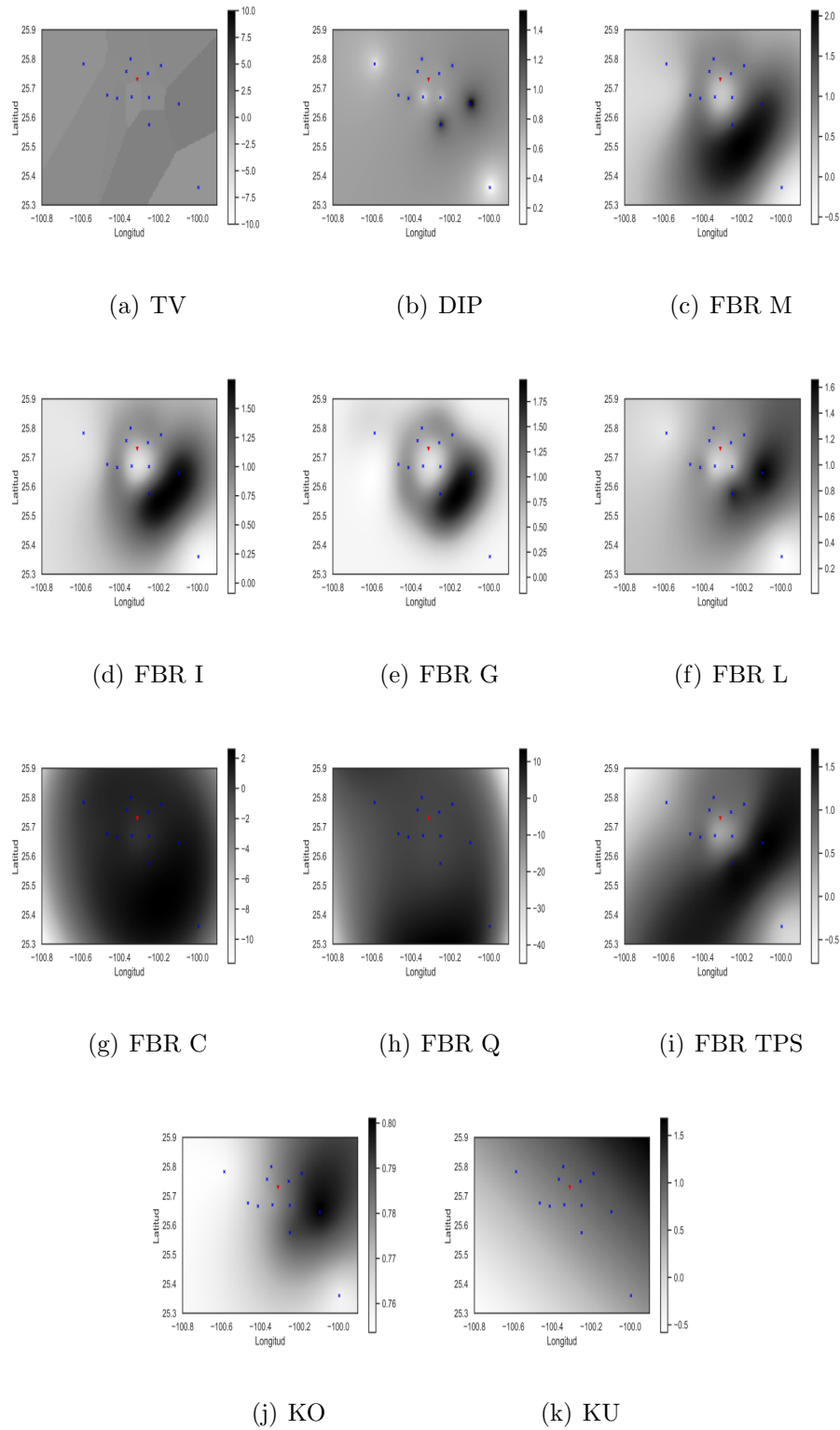


FIGURA 6.36: Interpolaciones de *velocidad del viento* para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.2.4 DIRECCIÓN DEL VIENTO

TABLA 6.37: *Dirección del viento*: 9 estaciones seleccionadas 4 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	3.69	72.45	0.59	104.70	0.86	10,962.57
DIP	3.13	54.17	0.44	79.79	0.65	6,367.46
FBR M	2.05×10^{16}	5.93×10^{17}	4.89×10^{15}	2.33×10^{18}	1.92×10^{16}	5.47×10^{36}
FBR IM	6.65×10^{15}	2.11×10^{17}	1.73×10^{15}	1.39×10^{18}	1.14×10^{16}	1.94×10^{36}
FBR G	2.09×10^{16}	5.41×10^{17}	4.46×10^{15}	2.23×10^{18}	1.84×10^{16}	4.99×10^{36}
FBR L	6.21×10^{11}	8.76×10^{13}	7.22×10^{11}	2.84×10^{16}	2.34×10^{14}	8.08×10^{32}
FBR C	3.73×10^{16}	1.52×10^{18}	1.25×10^{16}	3.75×10^{18}	3.09×10^{16}	1.40×10^{37}
FBR Q	8.66×10^{16}	2.60×10^{18}	2.14×10^{16}	4.89×10^{18}	4.03×10^{16}	2.39×10^{37}
FBR TPS	2.84×10^{14}	1.48×10^{17}	1.22×10^{15}	1.17×10^{18}	9.65×10^{15}	1.37×10^{36}
KO	3.11	55.24	0.45	81.71	0.67	6,677.18
KU	4.62×10^{15}	3.01×10^{17}	2.48×10^{15}	1.66×10^{18}	1.37×10^{16}	2.78×10^{36}

De la tabla 6.37, en la cual se utilizan nueve estaciones para interpolar cuatro estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son menores que los errores encontrados por el método KO, a excepción del MAPE. En la figura 6.37, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

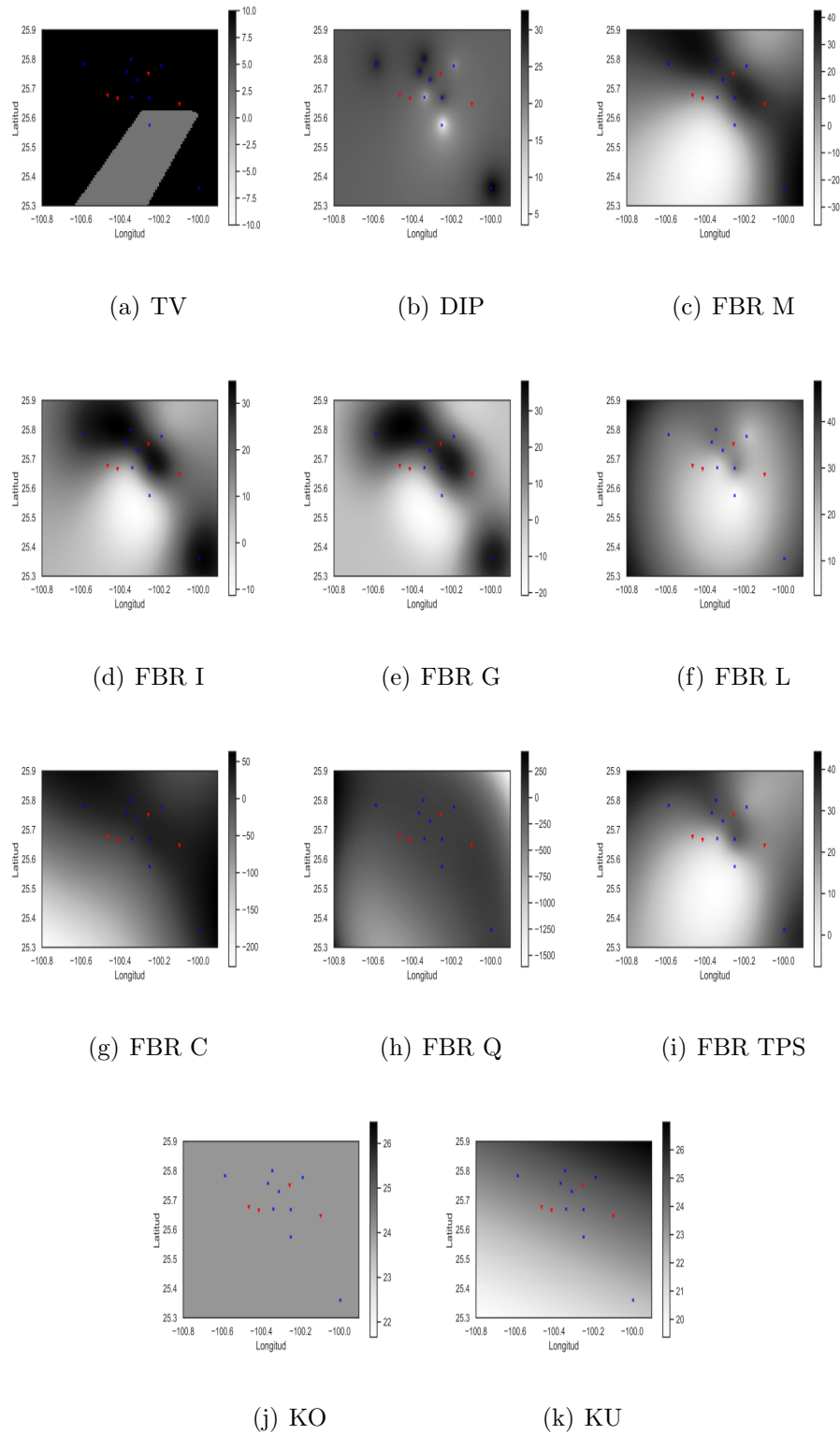


FIGURA 6.37: Interpolaciones de *dirección del viento* para 9 estaciones seleccionadas y 4 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.38: *Dirección del viento*: 10 estaciones seleccionadas 3 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	3.70	73.06	0.60	105.73	0.87	11,180.22
DIP	3.10	53.94	0.44	79.50	0.65	6,321.65
FBR M	2.16×10^{16}	6.91×10^{17}	5.69×10^{16}	2.52×10^{18}	2.08×10^{16}	6.37×10^{36}
FBR IM	7.05×10^{15}	2.39×10^{17}	1.97×10^{15}	1.48×10^{18}	1.22×10^{16}	2.21×10^{36}
FBR G	2.27×10^{16}	6.39×10^{17}	5.26×10^{15}	2.42×10^{18}	2.00×10^{16}	5.89×10^{36}
FBR L	3.82	83.42	0.68	134.67	1.10	18,137.55
FBR C	3.55×10^{16}	1.43×10^{18}	1.18×10^{16}	3.64×10^{18}	3.00×10^{16}	1.32×10^{37}
FBR Q	9.19×10^{16}	2.49×10^{18}	2.05×10^{16}	4.79×10^{18}	3.94×10^{16}	2.29×10^{37}
FBR TPS	3.16×10^{15}	1.59×10^{17}	1.31×10^{15}	1.21×10^{18}	9.98×10^{15}	1.47×10^{36}
KO	3.07	55.02	0.45	81.36	0.67	6,620.16
KU	4.55×10^{16}	3.04×10^{17}	2.50×10^{15}	1.67×10^{18}	1.37×10^{16}	2.80×10^{36}

De la tabla 6.38, en la cual se utilizan diez estaciones para interpolar tres estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son más pequeños que los errores de KO, a excepción del MAPE. En la figura 6.38, se pueden observar las interpolaciones de cada método, donde las puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

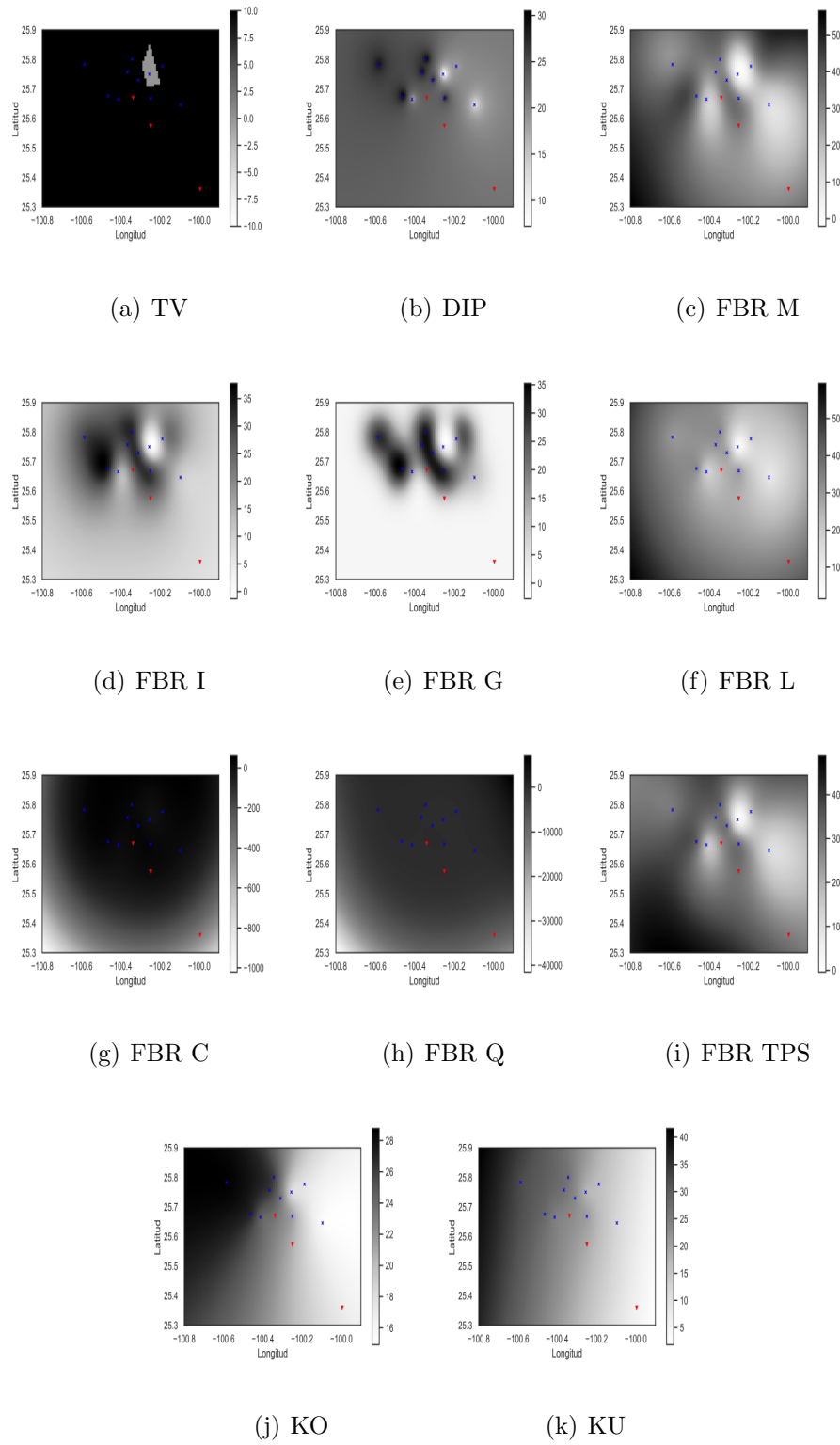


FIGURA 6.38: Interpolaciones de *dirección del viento* para 10 estaciones seleccionadas y 3 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.39: *Dirección del viento*: 11 estaciones seleccionadas 2 estaciones interpoladas

Métricas de error						
Método	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	3.63	72.71	0.60	105.58	0.87	11,148.52
DIP	3.05	53.14	0.43	78.74	0.65	6,200.05
FBR M	2.21×10^{16}	7.77×10^{17}	6.42×10^{14}	2.67×10^{18}	2.21×10^{16}	7.17×10^{36}
FBR IM	8.39×10^{15}	2.89×10^{17}	2.39×10^{15}	1.63×10^{18}	1.35×10^{16}	2.66×10^{36}
FBR G	2.40×10^{16}	7.86×10^{17}	6.50×10^{15}	2.69×10^{18}	2.22×10^{16}	7.25×10^{36}
FBR L	3.65	79.87	0.66	124.35	1.02	15,463.69
FBR C	3.33×10^{16}	1.38×10^{18}	1.14×10^{16}	3.57×10^{18}	2.95×10^{16}	1.27×10^{37}
FBR Q	8.77×10^{16}	41.99×10^{18}	1.64×10^{16}	4.28×10^{18}	3.54×10^{16}	1.83×10^{37}
FBR TPS	3.22×10^{15}	1.75×10^{17}	1.45×10^{15}	1.27×10^{18}	1.05×10^{16}	1.62×10^{36}
KO	3.04	54.07	0.44	80.26	0.66	6,442.88
KU	4.19×10^{15}	2.97×10^{17}	2.45×10^{15}	1.65×10^{18}	1.36×10^{16}	2.74×10^{36}

De la tabla 6.39, en la cual se utilizan once estaciones para interpolar dos estaciones, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE; mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO, pues en los errores MAPE, MAE, MAEP RMSE y MSE obtuvo mejores resultados. En la figura 6.39, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

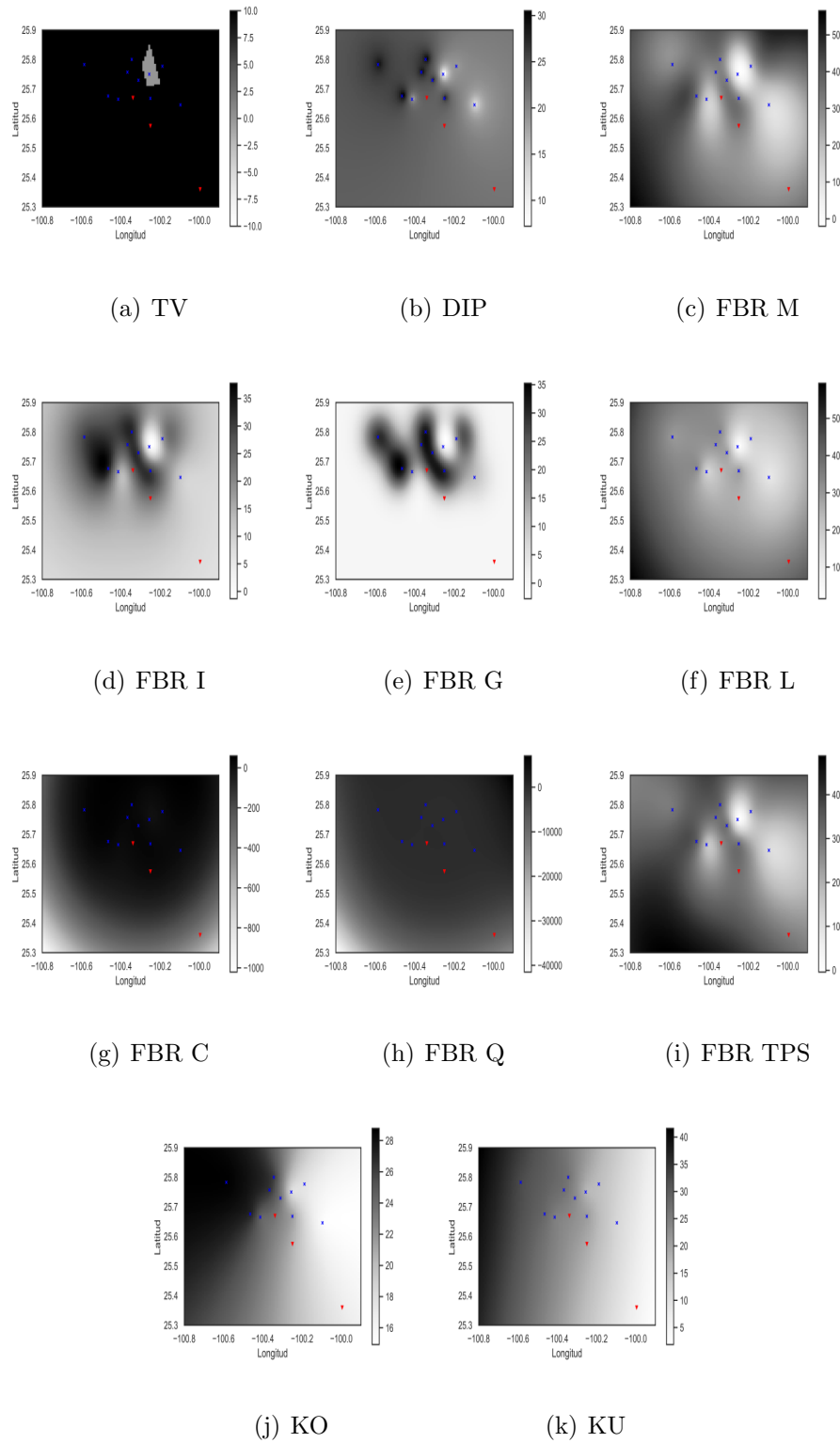


FIGURA 6.39: Interpolaciones de *dirección del viento* para 11 estaciones seleccionadas y 2 estaciones interpoladas: Fecha (31-12-2018 23:00:00)

TABLA 6.40: *Dirección del viento*: 12 estaciones seleccionadas 1 estaciones interpoladas

Método	Métricas de error					
	MAPE	MAE	MAEP	RMSE	RMSEP	MSE
TV	3.88	72.85	0.60	106.03	0.87	11,244.31
DIP	3.31	52.93	0.43	78.76	0.65	6,203.69
FBR M	2.31×10^{16}	8.69×10^{17}	7.20×10^{15}	2.83×10^{18}	2.34×10^{16}	8.02×10^{36}
FBR IM	5.98×10^{15}	3.21×10^{17}	2.66×10^{15}	1.72×10^{18}	1.42×10^{16}	2.96×10^{36}
FBR G	2.50×10^{16}	1.01×10^{18}	8.41×10^{15}	3.05×10^{18}	2.53×10^{16}	9.36×10^{36}
FBR L	4.01	77.75	0.64	118.30	0.98	13,996.62
FBR C	2.92×10^{16}	1.33×10^{18}	1.10×10^{16}	3.50×10^{18}	2.90×10^{16}	1.22×10^{37}
FBR Q	8.80×10^{16}	1.70×10^{18}	1.41×10^{16}	3.96×10^{18}	3.28×10^{16}	1.57×10^{37}
FBR TPS	3.49×10^{15}	1.76×10^{17}	1.45×10^{15}	1.27×10^{18}	1.05×10^{16}	1.62×10^{36}
KO	3.31	54.09	0.44	80.35	0.66	6,456.29
KU	3.52×10^{15}	2.82×10^{17}	2.34×10^{15}	1.61×10^{18}	1.33×10^{16}	2.60×10^{36}

De la tabla 6.40, en la cual se utilizan doce estaciones para interpolar una estación, podemos ver que los métodos que obtiene peores resultados de predicción son los métodos de Funciones de Base Radial, entre los métodos deterministas el método DIP obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, mientras que de los métodos geoestadísticos el método KO obtuvo el menor MAPE, MAE, MAEP, RMSE, RMSEP y MSE, en general DIP y KO son mejores que el resto de los métodos pero DIP es mejor que KO ya que en la mayoría de los errores son más pequeños que los errores de KO, a excepción del MAPE. En la figura 6.40, se pueden observar las interpolaciones de cada método, donde los puntos azules son las estaciones seleccionadas y los puntos rojos son las estaciones interpoladas.

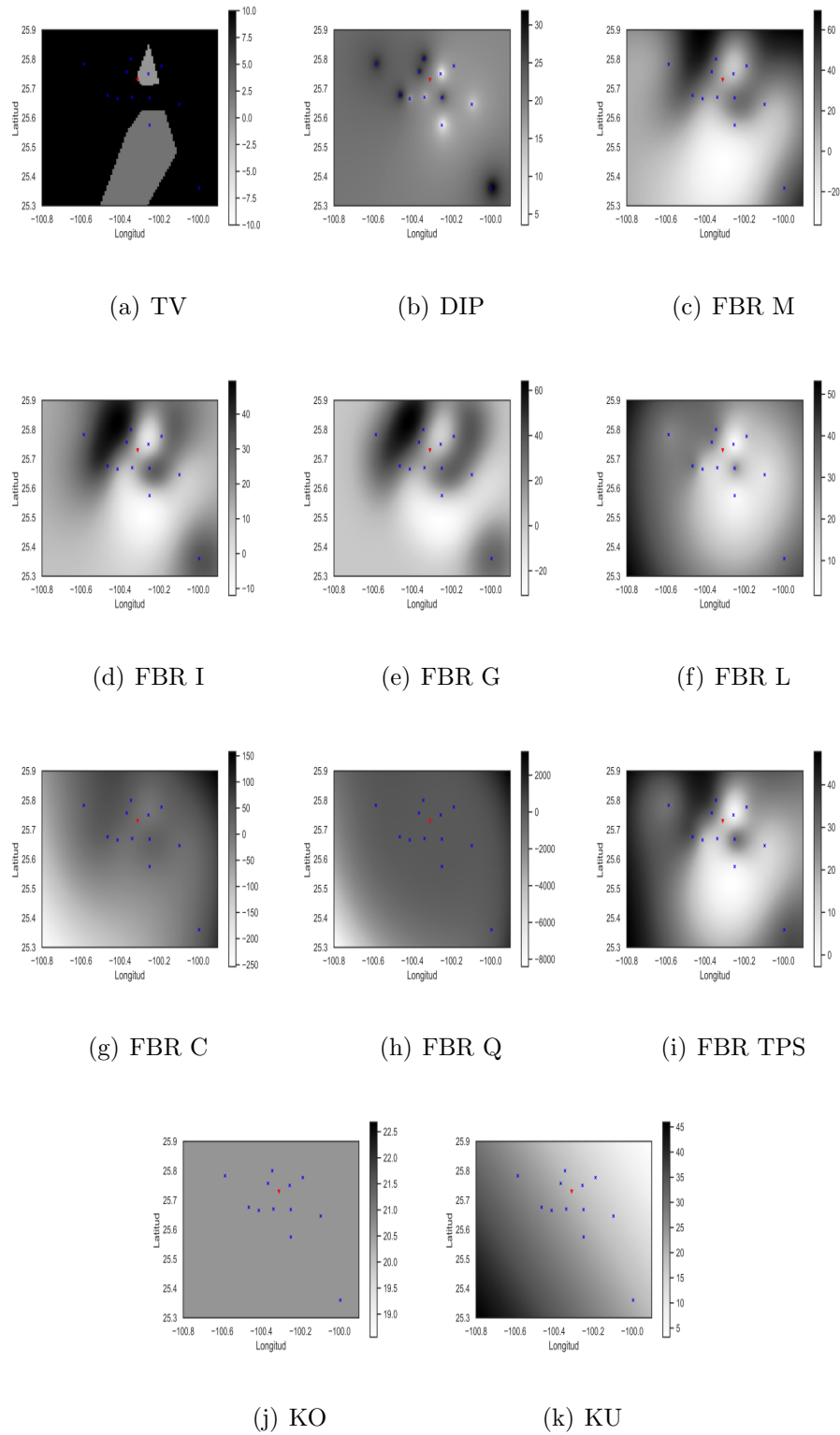


FIGURA 6.40: Interpolaciones de *dirección del viento* para 12 estaciones seleccionadas y 1 estación interpolada: Fecha (31-12-2018 23:00:00)

6.3. Conclusiones

Por último se presentan las conclusiones de los resultados obtenidos para cada una de las variables, esto es, se concluye cual es el mejor método de interpolar para cada variable, pues en todas las variables se encontraron buenos resultados, con diferentes métodos de interpolación por ejemplo, para las variables como: CO, PM₁₀, NO y NO_x el mejor interpolador es el método KO, mientras que para las variables NO₂, O₃, PM_{2.5}, SO₂, *velocidad del viento* y *dirección del viento* el mejor interpolador es el método DIP.

6.3.1 MONÓXIDO DE CARBONO (CO)

De la figura 6.41 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.4, es indiferente que error se seleccione ya que el método KO obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador para el contaminante CO es el método Kriging Ordinario** obteniendo un MAPE de 0.84, MAE de 0.88, MAEP de 0.66, RMSE de 1.11, RMSEP de 0.83 y MSE de 1.24.

6.3.2 BIÓXIDO DE NITRÓGENO (NO₂)

De la figura 6.42 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.8, es indiferente que error se seleccione ya que el método DIP obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador**

para el contaminante NO_2 es el método **Distancia Inversa Ponderada** obteniendo un MAPE de 0.59, MAE de 4.42, MAEP de 0.42, RMSE de 7.53, RMSEP de 0.72 y MSE de 56.79.

6.3.3 OZONO (O_3)

De la figura 6.43 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.12, es indiferente que error se seleccione ya que el método DIP obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador para el contaminante O_3 es el método Distancia Inversa Ponderada** obteniendo un MAPE de 0.48, MAE de 7.80, MAEP de 0.30, RMSE de 12.92, RMSEP de 0.50 y MSE de 167.16.

6.3.4 PARTÍCULAS MENORES A 10 MICRAS (PM_{10})

De la figura 6.44 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.16, es indiferente que error se seleccione ya que el método KO obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador para el contaminante PM_{10} es el método Kriging Ordinario** obteniendo un MAPE de 0.37, MAE de 19.03, MAEP de 0.31, RMSE de 33.51, RMSEP de 0.55 y MSE de 1,123.5.

6.3.5 PARTÍCULAS MENORES A 2.5 MICRAS ($PM_{2.5}$)

De la figura 6.45 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.20, es indiferente que error se seleccione ya que el método DIP obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador para el contaminante $PM_{2.5}$ es el método Distancia Inversa Ponderada** obteniendo un MAPE de 0.38, MAE de 6.50, MAEP de 0.29, RMSE de 11.90, RMSEP de 0.54 y MSE de 141.82.

6.3.6 BIÓXIDO DE AZUFRE (SO_2)

De la figura 6.46 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.24, es indiferente que error se seleccione ya que el método DIP obtiene los mejores resultados de interpolación, por lo que el **mejor interpolador para el contaminante SO_2 es el método Distancia Inversa Ponderada** obteniendo un MAPE de 0.44, MAE de 3.09, MAEP de 0.39, RMSE de 5.03, RMSEP de 0.64 y MSE de 25.33.

6.3.7 ÓXIDO NÍTRICO (NO)

De la figura 6.47 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.28, el método DIP obtiene el menor MAPE mientras que el método KO obtiene el menor MAE, MAEP, RMSE, RMSEP y MSE, por lo que el **mejor interpolador para el contaminante NO es el método Kriging**

Ordinario, ya que se escoje como métrica el RMSE obteniendo un MAPE de 1.42, MAE de 17.39, MAEP de 0.80, RMSE de 47.11, RMSEP de 2.18 y MSE de 2,219.88.

6.3.8 ÓXIDOS DE NITRÓGENO (NO_X)

De la figura 6.48 (c), podemos ver que estadísticamente la cantidad de valores atípicos respecto al total de las observaciones son pocas, por otro lado de la tabla 6.32, el método DIP obtiene el menor MAPE, MAE y MAEP, mientras que el método KO obtiene el menor RMSE, RMSEP y MSE, por lo que el **mejor interpolador para el contaminante NO_X es el método Kriging Ordinario**, ya que se escoje como métrica el RMSE obteniendo un MAPE de 0.81, MAE de 25.90, MAEP de 0.66, RMSE de 54.61, RMSEP de 1.39 y MSE de 2,982.28.

6.3.9 VELOCIDAD DEL VIENTO

Para esta variable no se presentan imágenes de la interpolación ya que la velocidad del viento se suele representar por puntos espaciales a través de rosas de viento, por otro lado de la tabla 6.36, el método DIP obtiene el menor MAPE, MAE y MAEP, RMSE, RMSEP y MSE, por lo que el **mejor interpolador para la variable *velocidad del viento* es el método Distancia Inversa Ponderada**, obteniendo un MAPE de 0.52, MAE de 2.62, MAEP de 0.34, RMSE de 3.69, RMSEP de 0.48 y MSE de 13.64.

6.3.10 DIRECCIÓN DEL VIENTO

Para esta variable no se presentan imágenes de la interpolación ya que la velocidad del viento se suele representar por puntos espaciales a través de rosas de viento, por otro lado de la tabla 6.40, el método DIP obtiene el menor MAPE, MAE y MAEP, RMSE, RMSEP y MSE, por lo que el **mejor interpolador para la variable *dirección del viento* es el método Distancia Inversa Ponderada**, obteniendo un MAPE de 3.31, MAE de 52.93, MAEP de 0.43, RMSE de 78.76, RMSEP de 0.65 y MSE de 6,203.69.

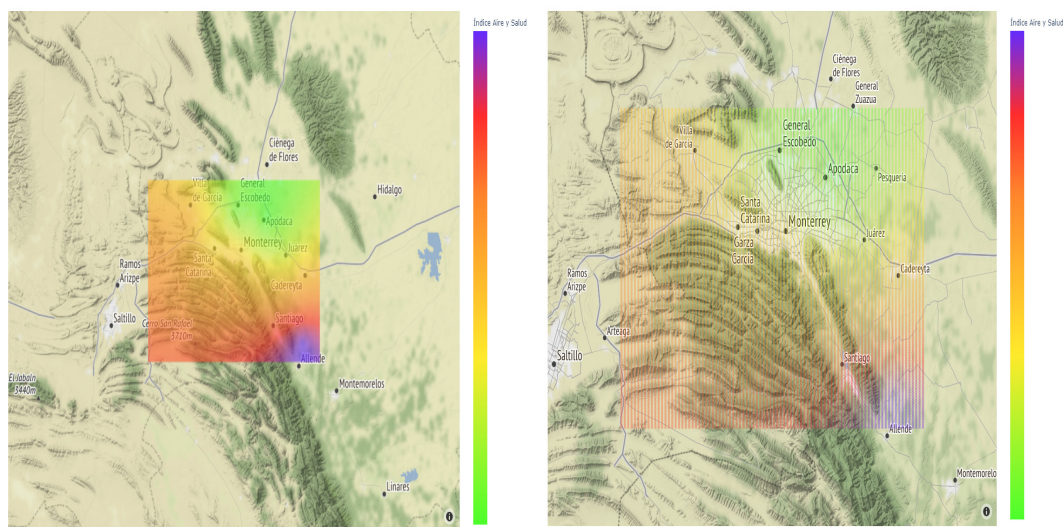
Todas las variables presentan resultados para los métodos buenos si son variables de contaminación en el caso de las variables de velocidad del viento y *dirección del viento* los métodos de interpolación no resultan ser tan buenos predictores. De las once métricas propuestas para interpolar, los métodos Kriging Ordinario y Distancia Inversa Ponderada resultan ser los mejores para interpolar ya que para las variables como: CO, PM₁₀, NO y NO_x el mejor interpolador es el método KO, mientras que para las variables NO₂, O₃, PM_{2.5}, SO₂, *velocidad del viento* y *dirección del viento* el mejor interpolador es el método DIP.

6.4. Trabajo Futuro

- Utilizar los resultados obtenidos para proponer una publicación en la cual se presenten los resultados obtenidos de la comparación hecha entre los métodos de interpolación espacial usados en este trabajo.
- Analizar relaciones entre datos multivariados, es decir, de las correlaciones cruzadas calculadas utilizar los valores obtenidos de correlaciones positivamente altas para ayudar a una variable a predecir a partir de otras correlaciones a ésta, por ejemplo utilizar la variable *radiación solar* para interpolar la variable

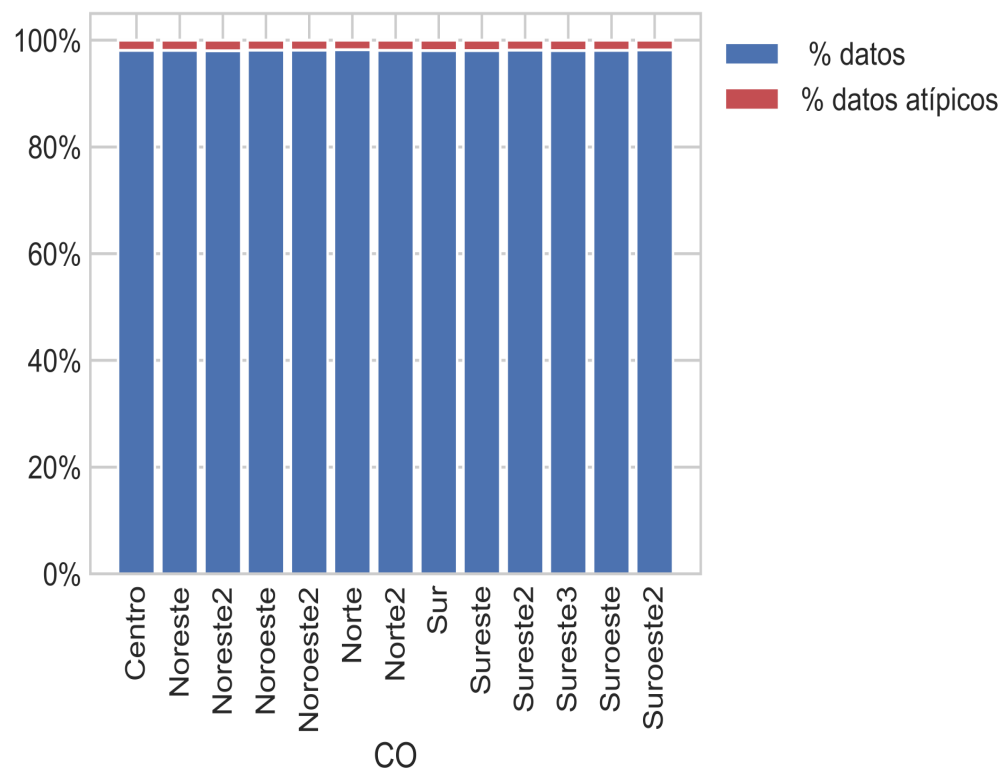
temperatura ambiental.

- Utilizar las autocorrelaciones calculadas para integrar la variable tiempo en las interpolaciones, es decir, de las autocorrelaciones encontradas utilizar para pronosticar a valor futuro, esto es utilizar los valores al tiempo t para pronosticar al tiempo $t + n$ donde la autocorrelación se maximice.
- Comparar con datos similares de otras ciudades, donde se hayan aplicado los métodos de interpolación implementados en este trabajo.
- Incorporar la variable tiempo a un modelo para que el modelo estadístico espacial ahora sea estadístico espacio-temporal.



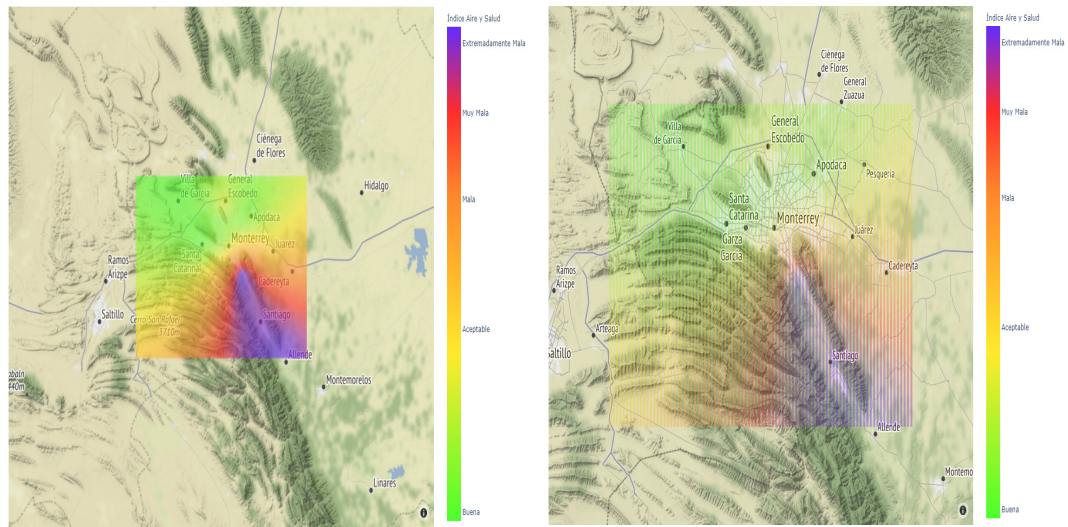
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



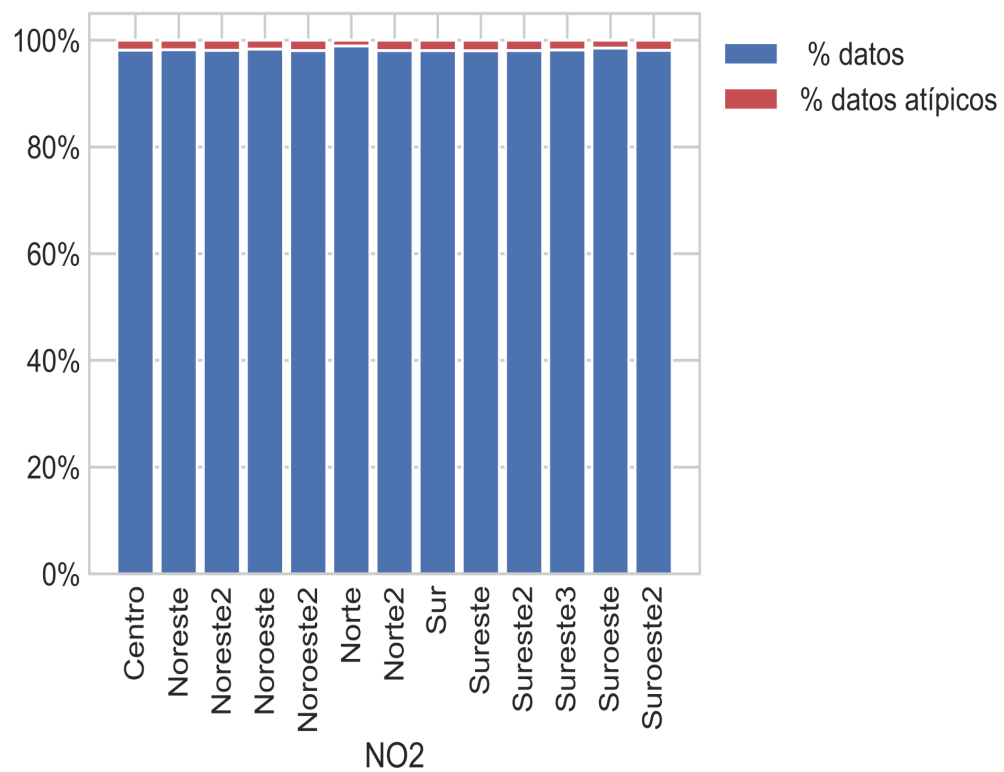
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.41: Interpolaciones de CO



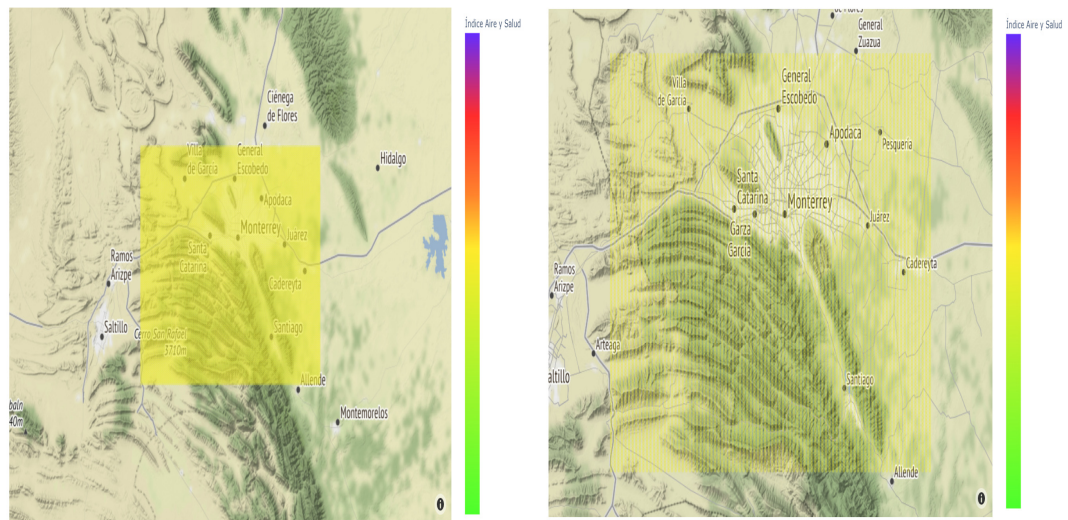
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



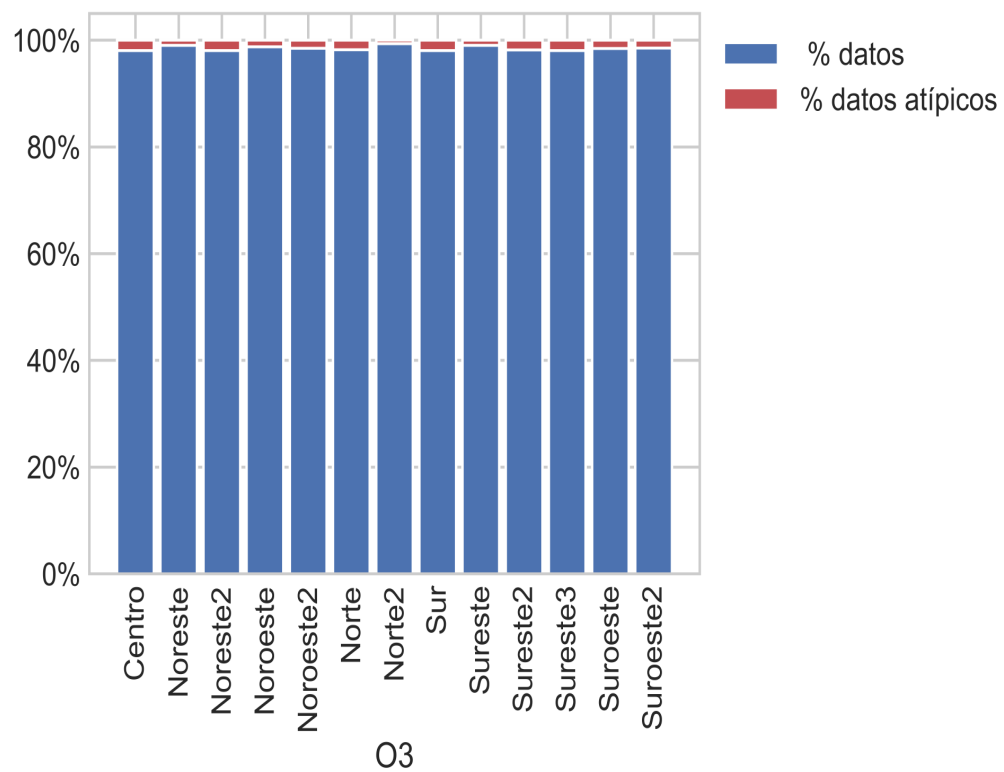
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.42: Interpolaciones de NO₂



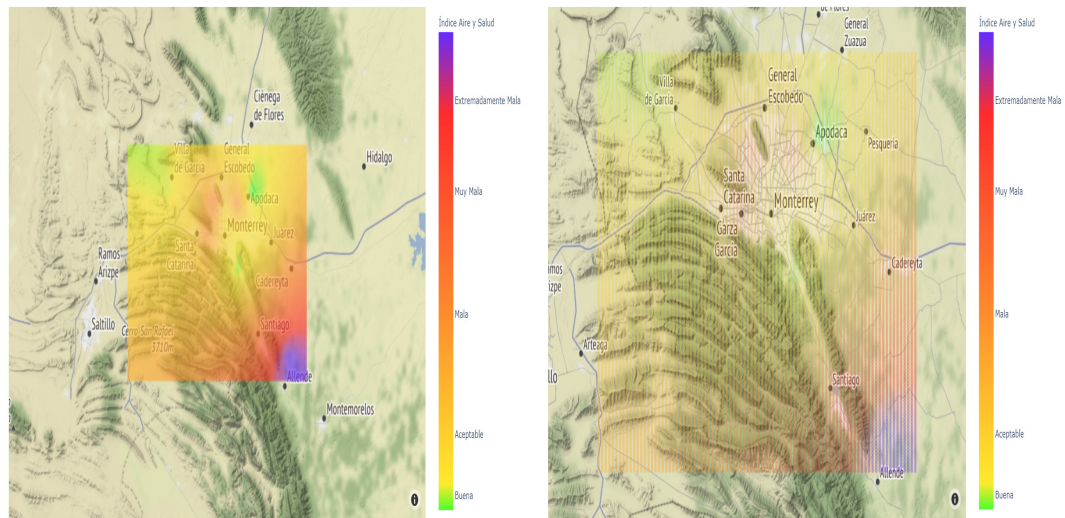
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



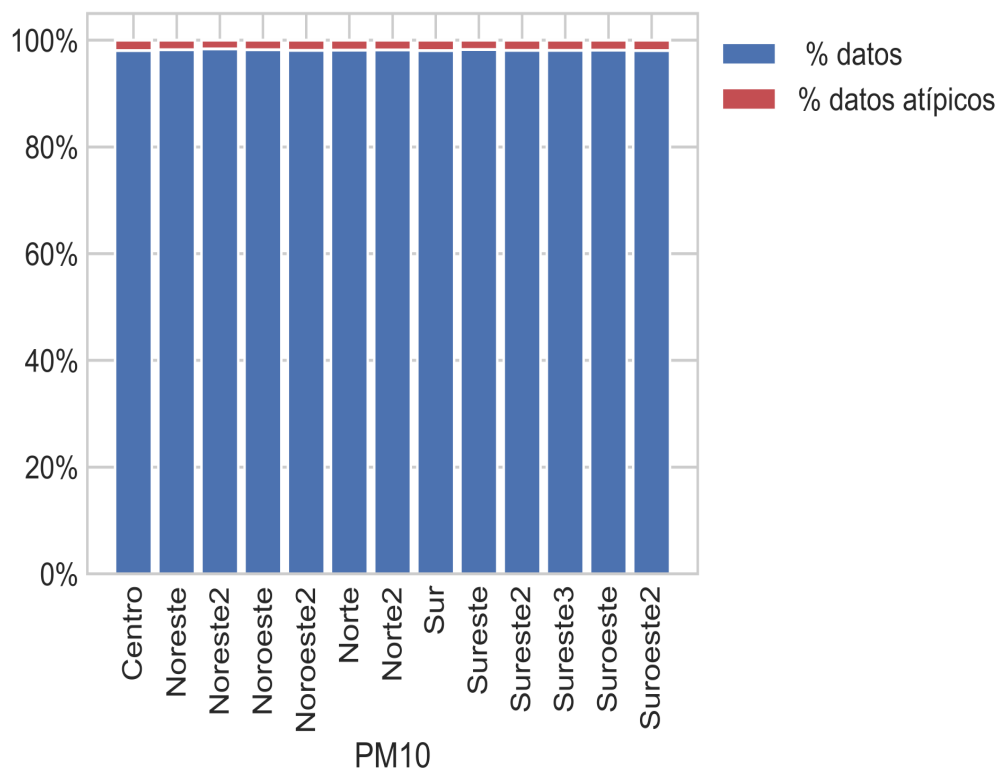
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.43: Interpolaciones de O_3



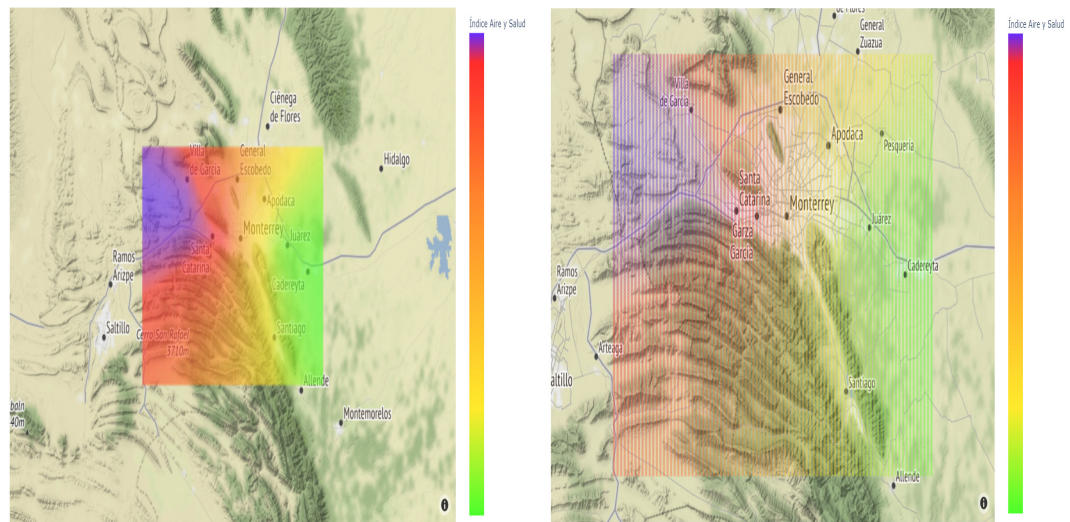
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



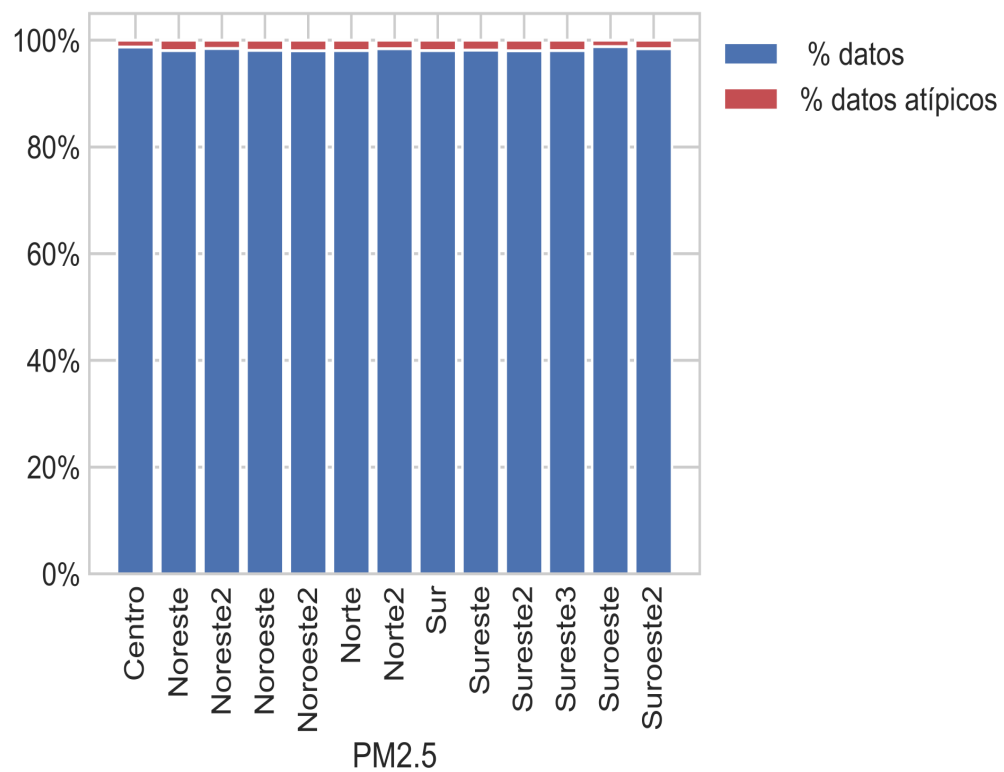
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.44: Interpolaciones de PM₁₀



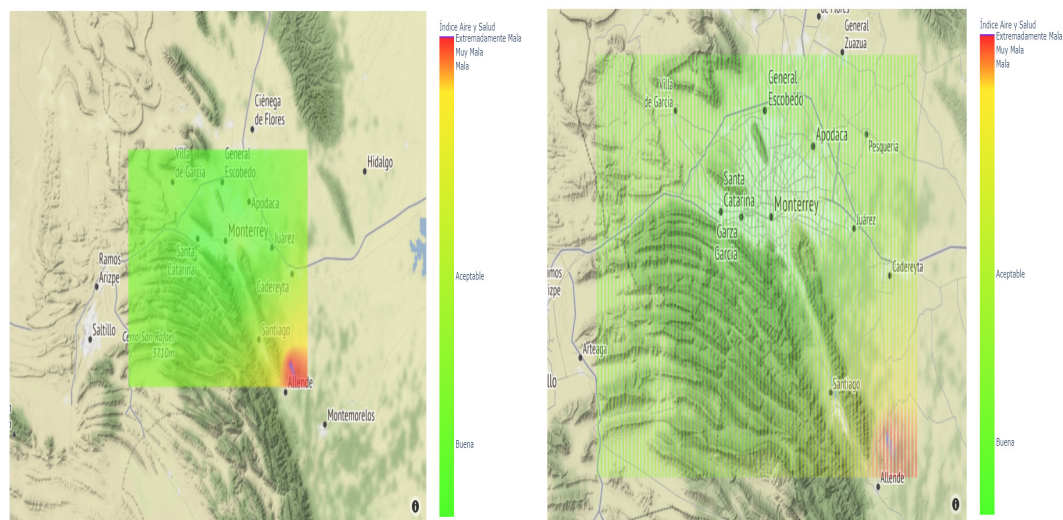
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



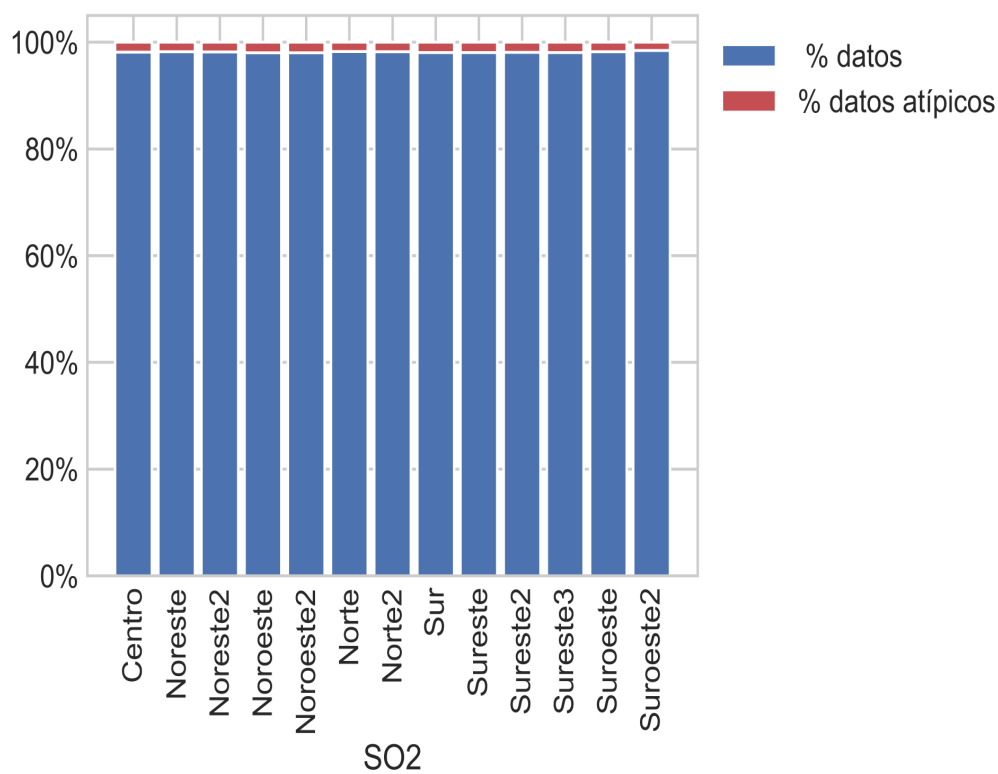
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.45: Interpolaciones de $PM_{2.5}$



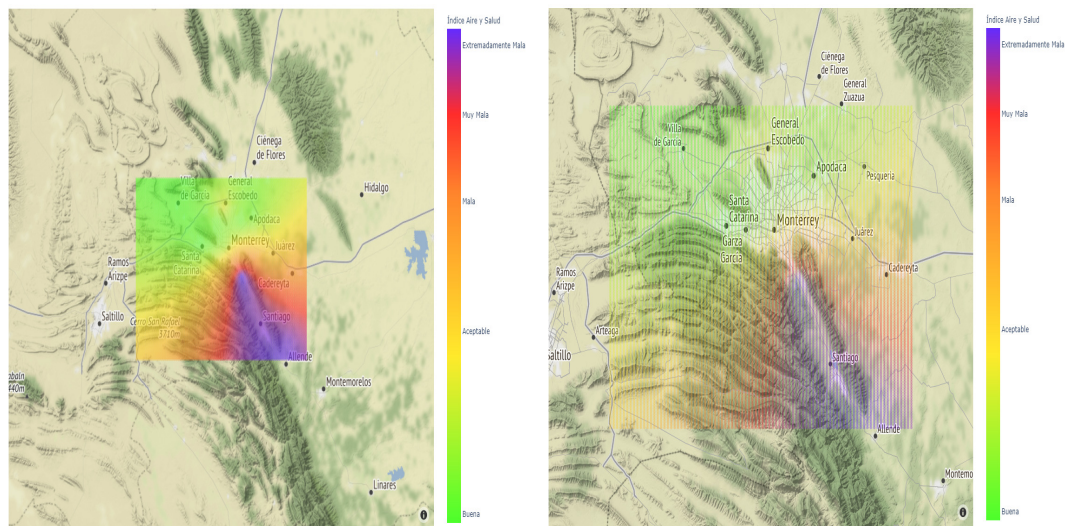
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



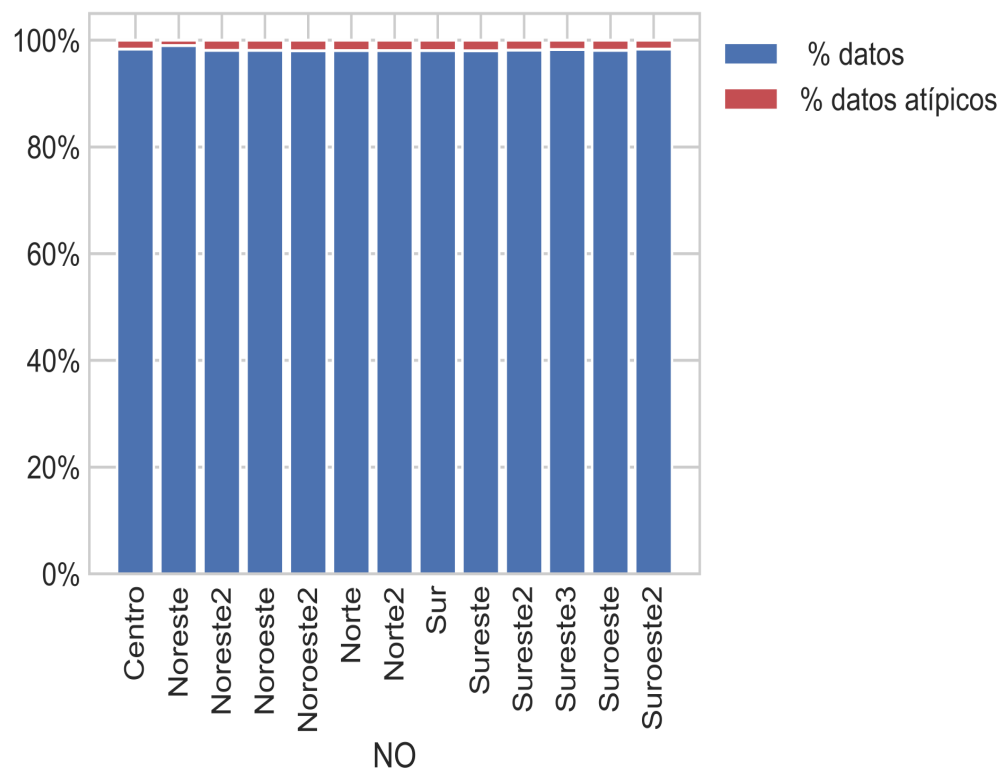
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.46: Interpolaciones de SO₂



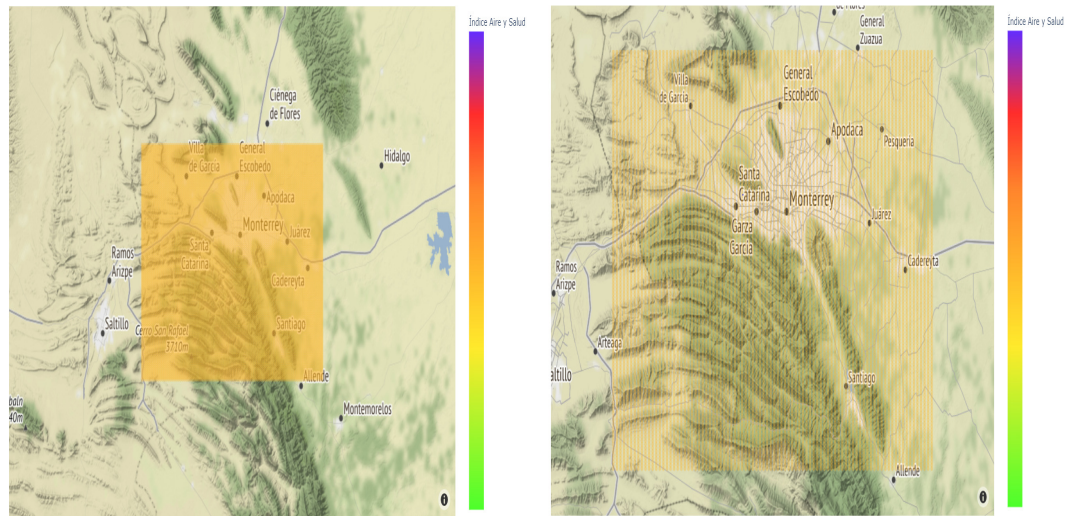
(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



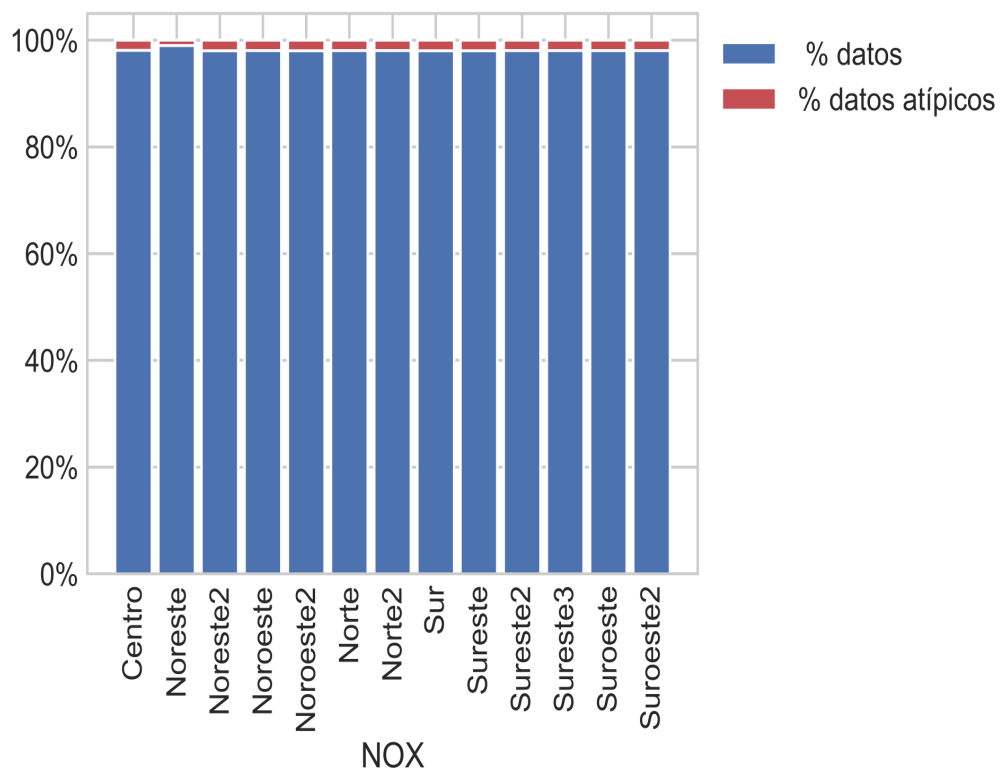
(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.47: Interpolaciones de NO



(a) Interpolación de las 13 estaciones con el método KO

(b) Interpolación con el método KO



(c) Porcentaje de datos y porcentaje de atípicos

FIGURA 6.48: Interpolaciones de NO_x

BIBLIOGRAFÍA

- AGUILAR, F. J., F. AGÜERA, M. A. AGUILAR y F. CARVAJAL (2005), «Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy», *Photogrammetric Engineering & Remote Sensing*, **71**(7), págs. 805–816.
- ALLOWAY, B. y D. C. AYRES (1997), *Chemical principles of environmental pollution*, CRC press.
- ARIAS, J. A., B. LOSADA, F. MORENO y A. BRAVO (2009), «Mejoramiento de imágenes usando funciones de base radial», *Revista Ingenierías Universidad de Medellín*, **8**(15), págs. 21–28.
- ARMSTRONG, M. P. y R. MARCIANO (1994), «Inverse-distance-weighted spatial interpolation using parallel supercomputers», *Photogrammetric Engineering and Remote Sensing*, **60**(9), págs. 1097–1102.
- AURENHAMMER, F. y R. KLEIN (2000), «Voronoi diagrams», *Handbook of computational geometry*, **5**(10), págs. 201–290.
- BABAK, O. y C. V. DEUTSCH (2009), «Statistical approach to inverse distance interpolation», *Stochastic Environmental Research and Risk Assessment*, **23**(5), págs. 543–553.
- BARTIER, P. M. y C. P. KELLER (1996), «Multivariate interpolation to incorporate

- thematic surface data using inverse distance weighting (IDW)», *Computers & Geosciences*, **22**(7), págs. 795–799.
- BAYRAKTAR, H. y F. S. TURALIOGLU (2005), «A Kriging-based approach for locating a sampling site—in the assessment of air quality», *Stochastic Environmental Research and Risk Assessment*, **19**(4), págs. 301–305.
- BEL, L. (2004), «Non Parametric Variogram Estimator. Application to Air Pollution Data», en *geoENV IV—Geostatistics for Environmental Applications*, Springer, págs. 29–40.
- BILONICK, R. A. (1985), «The space-time distribution of sulfate deposition in the Northeastern United States», *Atmospheric Environment (1967)*, **19**(11), págs. 1829–1845.
- BILONICK, R. A. (1988), «Monthly hydrogen ion deposition maps for the Northeastern US from July 1982 to September 1984», *Atmospheric Environment (1967)*, **22**(9), págs. 1909–1924.
- BROWN, G. S. (1965), *Point density in stems per acre*, Forest Research Institute, New Zealand Forest Service.
- BUXTON, B. E. y A. D. PATE (1994), «Joint temporal-spatial modeling of concentrations of hazardous pollutants in urban air», en *Geostatistics for the next century*, Springer, págs. 75–87.
- CHEN, F.-W. y C.-W. LIU (2012), «Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan», *Paddy and Water Environment*, **10**(3), págs. 209–222.
- CRESSIE, N. (1990), «The origins of kriging», *Mathematical geology*, **22**(3), págs. 239–252.
- CURRIE, J. y M. NEIDELL (2005), «Air pollution and infant health: what can we learn from California’s recent experience?», *The Quarterly Journal of Economics*, **120**(3), págs. 1003–1030.

- DELGADO, O. y J. MARTÍNEZ (2015), «Elaboración del mapa de ruido del área urbana de la Ciudad de Cuenca–Ecuador, empleando la técnica de interpolación geoestadística Kriging ordinario», *Ciencias Espaciales*, **8**(1), págs. 411–440.
- DIMITRAKOPOULOS, R. y X. LUO (1994), «Spatiotemporal modelling: covariances and ordinary kriging systems», en *Geostatistics for the next century*, Springer, págs. 88–93.
- DINGMAN, S. L., D. M. SEELY-REYNOLDS y R. C. REYNOLDS III (1988), «Application of kriging to estimating mean annual precipitation in a region of orographic influence 1», *JAWRA Journal of the American Water Resources Association*, **24**(2), págs. 329–339.
- DIRICHLET, G. L. (1850), «Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen», *J. reine angew. Math.*, **40**, pág. 209.
- DOVAL, H. C. (2003), «John Snow y la epidemia de cólera en Londres en 1854: “los hechos hablan por sí solos, o es necesaria una hipótesis para encontrar los hechos” [carta del director]», *Revista Argentina de Cardiología*, págs. 463–467.
- FENGER, J. (1999), «Urban air quality», *Atmospheric environment*, **33**(29), págs. 4877–4900.
- FRANKE, R. (1982), «Scattered data interpolation: tests of some methods», *Mathematics of computation*, **38**(157), págs. 181–200.
- GIROSI, F. (1992), «Some extensions of radial basis functions and their applications in artificial intelligence», *Computers & Mathematics with Applications*, **24**(12), págs. 61–80.
- GREENPEACE (2011), «No Apto Para Pulmones Pequeños, Diagnóstico de Calidad del Aire y el Derecho de Niñas, Niños y Adolescentes al Aire Limpio», https://www.greenpeace.org/archive-mexico/Global/mexico/Docs/2018/Revoluci%C3%B3n%20Urbana/No_apto_para_pulmones_pequen%C3%83osok3.pdf. Accedido: 23-01-2019.

- GREENPEACE (2019), «Demandan ONG acciones contundentes para mejorar la calidad del aire en Monterrey», <https://www.greenpeace.org/mexico/noticia/1648/demandan/ong/acciones/contundentes/para/mejorar/la/calidad/del/aire/en/monterrey/>. Accedido: 01-02-2019.
- HARDY, R. L. (1971), «Multiquadric equations of topography and other irregular surfaces», *Journal of geophysical research*, **76**(8), págs. 1905–1915.
- HARDY, R. L. (1990), «Theory and applications of the multiquadric-biharmonic method 20 years of discovery 1968–1988», *Computers & Mathematics with Applications*, **19**(8-9), págs. 163–208.
- HOUGH, R. L., N. BREWARD, S. D. YOUNG, N. M. CROUT, A. M. TYE, A. M. MOIR y I. THORNTON (2004), «Assessing potential risk of heavy metal exposure from consumption of home-produced vegetables by urban populations.», *Environmental health perspectives*, **112**(2), págs. 215–221.
- ICKE, V., R. WEYGAERT *et al.* (1987), «Fragmenting the universe», *Astronomy and Astrophysics*, **184**, pág. 16.
- KANAROGLU, P. S., N. SOULAKELLIS y N. SIFAKIS (2002), «Improvement of satellite derived pollution maps with the use of a geostatistical interpolation method», *Journal of Geographical Systems*, **4**(2), págs. 193–208.
- KIŠ, I. M. (2016), «Comparison of Ordinary and Universal Kriging interpolation techniques on a depth variable (a case of linear spatial trend), case study of the Šandrovac Field», *Rudarsko-geološko-naftni zbornik*, **31**(2), págs. 41–58.
- KOLMOGOROV, A. (1941), «Interpolirovanie i ekstrapolirovanie statsionarnykh sluchainykh posledovatel'nostei», *Izvestiya Akademii Nauk Sssr Seriya Fizicheskaya*, **5**, págs. 3–14.
- KOLMOGOROV, A. N. (1939), «Sur l'interpolation et extrapolation des suites stationnaires», *Comptes rendus de l'Académie des Sciences*, **208**, págs. 2043–2045.

- KREMPER, A., T. SCHANZE y R. ECKHORN (2002), «Classification of neural signals by a generalized correlation classifier based on radial basis functions», *Journal of neuroscience methods*, **116**(2), págs. 179–187.
- KUMAR, V. (2007), «Optimal contour mapping of groundwater levels using universal kriging—a case study», *Hydrological Sciences Journal*, **52**(5), págs. 1038–1050.
- LAZZARO, D. y L. B. MONTEFUSCO (2002), «Radial basis functions for the multivariate interpolation of large scattered data sets», *Journal of Computational and Applied Mathematics*, **140**(1-2), págs. 521–536.
- MANZOOR, S., M. EDWARDS, A. DOGRU y T. AL-SHAALAN (2014), «Boundary Aligned Grid Generation and CVD-MPFA-Cell-centred Versus Cell-vertex on Unstructured Grids», en *ECMOR XIV-14th European conference on the mathematics of oil recovery*, tomo 2014, European Association of Geoscientists & Engineers, págs. 1–30.
- MARTÍNEZ, J., N. ARELLANO y S. SÁENZ (2001), «Análisis de las Mediciones de la Calidad del Aire en el Área Metropolitana de Monterrey: SIMA (1993-2001)», *Análisis de Riesgo: Una herramienta de prevención en la planeación urbana*, págs. 8–13.
- MATHERON, G. (1963), «Principles of geostatistics», *Economic geology*, **58**(8), págs. 1246–1266.
- MATHERON, G. (1965), *Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature*, Masson et CIE.
- MAYER, H. (1999), «Air pollution in cities», *Atmospheric environment*, **33**(24-25), págs. 4029–4037.
- MERCER, L. D., A. A. SZPIRO, L. SHEPPARD, J. LINDSTRÖM, S. D. ADAR, R. W. ALLEN, E. L. AVOL, A. P. ORON, T. LARSON, L.-J. S. LIU *et al.* (2011), «Comparing universal kriging and land-use regression for predicting concentrations

- of gaseous oxides of nitrogen (NO_x) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air)», *Atmospheric Environment*, **45**(26), págs. 4412–4420.
- MONTGOMERY, D. C., C. L. JENNINGS y M. KULAHCI (2015), *Introduction to time series analysis and forecasting*, John Wiley & Sons. ISBN 978-0-471-65397-4.
- OLIVER, M. A. y R. WEBSTER (1990), «Kriging: a method of interpolation for geographical information systems», *International Journal of Geographical Information System*, **4**(3), págs. 313–332.
- OLIVER, M. A. y R. WEBSTER (2015), *Basic steps in geostatistics: the variogram and kriging*, Springer.
- ORGANIZACIÓN MUNDIAL DE LA SALUD, O. M. S. (2019), «Contaminación del Aire Ambiental», https://www.paho.org/hq/index.php?option=com_content&view=article&id=12918:ambient-air-pollution&Itemid=72243&lang=es.
Accedido: 16-09-2019.
- SAMPER, F. y J. CARRERA (1990), «Geoestadística», *Aplicaciones a la Hidrogeología Subterránea. Centro Internacional de Métodos Numéricos en Ingeniería. Universitat Politècnica de Catalunya. Barcelona*.
- SAMPSON, P. D., M. RICHARDS, A. A. SZPIRO, S. BERGEN, L. SHEPPARD, T. V. LARSON y J. D. KAUFMAN (2013), «A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM_{2.5} concentrations in epidemiology», *Atmospheric environment*, **75**, págs. 383–392.
- SECRETARÍA DE DESARROLLO SUSTENTABLE DE NUEVO LEÓN, N. L. (2019), «Reporte de Calidad del Aire y Meteorología del Área Metropolitana de Monterrey», http://aire.nl.gob.mx/docs/reportes/mensuales/12_Reporte_Diciembre_2018.pdf. Accedido: 01-04-2019.

- SECRETARÍA DE DESARROLLO SUSTENTABLE DE NUEVO LEÓN, N. L. (2020), «Índice Aire y Salud. Reporte de la calidad del aire en la Zona Metropolitana de Monterrey», http://aire.nl.gob.mx/icars2020/map_calidad_icars.php. Accedido: 23-04-2020.
- SENECHAL, M. (1993), «Spatial tessellations: Concepts and applications of voronoi diagrams», *Science*, **260**(5111), págs. 1170–1173.
- SERNA, A. (2020), «Modelos estadístico-espaciales de contaminantes del aire en el Área Metropolitana de Monterrey», <https://github.com/SernArmando/Air-pollution-Monterrey>. Accedido: 07-09-2020.
- SHEPARD, D. (1968), «A two-dimensional interpolation function for irregularly-spaced data», en *Proceedings of the 1968 23rd ACM national conference*, págs. 517–524.
- SONAR, T. (1996), «Optimal recovery using thin plate splines in finite volume methods for the numerical solution of hyperbolic conservation laws», *IMA Journal of Numerical Analysis*, **16**(4), págs. 549–581.
- TALMI, A. y G. GILAT (1977), «Method for smooth approximation of data», *Journal of Computational Physics*, **23**(2), págs. 93–123.
- TANG, M., P. AGRAWAL, F. NIE, S. PONGPAICHET y R. JAIN (2016), «A graph based multimodal geospatial interpolation framework», en *2016 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE, págs. 1–6.
- TUREK, Z., L. HOOFD, S. BATRA y K. RAKUSAN (1992), «The effect of realistic geometry of capillary networks on tissue PO_2 in hypertrophied rat heart», en *Oxygen Transport to Tissue XIV*, Springer, págs. 567–572.
- VILLATORO, M., C. R. HENRÍQUEZ HENRÍQUEZ y F. SANCHO (2007), «Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, Ca, CICE y P del suelo», *Agronomía Costarricense Vol. 32 Núm. 1 2007*.

- VOGL, S., P. LAUX, W. QIU, G. MAO y H. KUNSTMANN (2012), «Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields», *Hydrology and Earth System Sciences*, **16**, págs. 2311–2328.
- WARE, C., W. KNIGHT y D. WELLS (1991), «Memory intensive statistical algorithms for multibeam bathymetric data», *Computers & Geosciences*, **17**(7), págs. 985–993.
- WEBSTER, R. y M. A. OLIVER (2007), *Geostatistics for environmental scientists*, John Wiley & Sons. ISBN 978-047-002-858-2.
- WIENER, N. (1964), *Extrapolation, interpolation, and smoothing of stationary time series*, The MIT press. ISBN 978-026-225-719-0.
- WOLD, H. (1938), *A study in the analysis of stationary time series*, Tesis Doctoral, Stockholm College.
- XIE, Y., T.-B. CHEN, M. LEI, J. YANG, Q.-J. GUO, B. SONG y X.-Y. ZHOU (2011), «Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis», *Chemosphere*, **82**(3), págs. 468–476.

RESUMEN AUTOBIOGRÁFICO

Jorge Armando Serna Mendoza

Candidato para obtener el grado de
Maestría en Ciencias de la Ingeniería
con Orientación en Sistemas

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica

Tesis:

MODELOS ESTADÍSTICO-ESPACIALES DE CONTAMINANTES DEL
AIRE EN EL ÁREA METROPOLITANA DE MONTERREY

Nací el 1 de septiembre de 1995 en la Ciudad de México, México; mis padres son Armando Serna Suárez y Alma Delia Mendoza Cabello. En 2019 egresé como Licenciado en Actuaría en la Facultad de Ciencias de la Universidad Nacional Autónoma de México (UNAM).